

HÅNDBOK

2018

METODEHÅNDBOK

Slik oppsummerer vi
forskning

Tittel	Slik oppsummerer vi forskning. Håndbok for Folkehelseinstituttet.
Institusjon	Folkehelseinstituttet
Ansvarlig	Camilla Stoltenberg, <i>direktør</i>
ISBN	978-82-8121-427-9
Antall sider	71 (85 med vedlegg)
Oppdragsgiver	Område for helsetjenester i Folkehelseinstituttet
Sitering	Område for helsetjenester i Folkehelseinstituttet. Slik oppsummerer vi forskning. Håndbok for Folkehelseinstituttet. 4. reviderte utg. Oslo: Område for helsetjenester i Folkehelseinstituttet; 2018.

Område for helsetjenester i Folkehelseinstituttet leverer kunnskapsgrunnlag til beslutningstakere på alle nivåer i helse-, omsorgs- og velferdstjenestene, fra sentrale myndigheter til kommunehelsetjenesten. Målet er å bidra til gode beslutninger slik at brukerne får best mulig tjenester. Folkehelseinstituttet kan ikke instrueres i faglige spørsmål.

Område for helsetjenester i Folkehelseinstituttet
Oslo, august 2018

Innholdsfortegnelse

INNHALDSFORTEGNELSE	2
FORORD	4
1 VÅRE ULIKE OVERSIKTER	5
Systematisk oversikt	7
Oversikt over systematiske oversikter	8
Hurtigoversikt	9
Metodevurdering	9
Metodevarsel	10
Systematisk kartleggingsoversikt	10
Systematisk litteratursøk med sortering	11
Omtale av systematisk oversikt	12
2 OVERBLIKK: SLIK OPPSUMERER VI FORSKNING	13
Hvem gjør hva i prosessen?	14
3 FORMULERE SPØRSMÅL OG SKRIVE PROSJEKTPLAN	16
Formulere spørsmål	17
Skrive prosjektplan	22
4 SØKE ETTER LITTERATUR	26
Hva kjennetegner gode litteratursøk?	26
Velge ut databaser og andre litteraturkilder	27
Lage søkealgoritme	30
Vurdere resultatene og eventuelt revidere søkestrategiene	33
Bearbeide referansene ved hjelp av referansehåndteringsverktøy	33
Loggføre og rapportere søk	33
Oppdatere søk	33
5 VELGE UT STUDIER	35
Er artikkelsammendraget relevant?	36
Er selve publikasjonen relevant?	36
6 VURDERING AV INKLUDERTE STUDIER	38
Vurdering av inkluderte studier	39
Å vurdere oversiktens metodiske kvalitet	40
Å vurdere systematiske skjevheter i primærstudier	40

Samlet vurdering av risiko for systematiske skjevheter	41
7 HENTE UT DATA, SAMMENSTILLE OG GRADERE	43
Hente ut data	43
Bruk av data fra systematiske oversikter	45
Bruk av data fra primærstudier	45
Beskrivelse av studiene og resultatene	46
Gradering	49
8 SKRIVE OVERSIKTEN	54
Innholdet i rapportene	55
9 FAGFELLEVRDERE, GODKJENNE OG PUBLISERE	58
Fagfellevurdering	59
Publisere oversiktene	59
10 DIAGNOSTIKK	60
11 KVALITATIVE OPPSUMMERINGER	61
12 METODEVURDERINGER	62
Innledning	62
Helseøkonomi	63
Etiske vurderinger det er foreslått å bytte ut moral og moralske med etisk og etiske, ok?	68
Juridiske betraktninger, Organisatoriske betraktninger, Andre betraktninger?	69
LITTERATUR	71

Forord

Håndboken «Slik oppsummerer vi forskning» er viktig for ansatte i område for helsetjenester og andre ved Folkehelseinstituttet som vurderer og oppsummerer forskningsbasert kunnskap. Vi ønsker å produsere systematiske oversikter (også kalt systematiske forskningsoversikter eller kunnskapsoppsummeringer) av god metodisk kvalitet som er pålitelige og har relevans og er til nytte for helse-, omsorgs- og velferdstjenestene. Målet er at oversiktene skal bidra til bedre beslutninger og økt kvalitet på alle nivåer i tjenestene.

Håndboken viser våre ulike produkter og framgangsmåtene vi benytter. Vi håper at også andre enn våre ansatte, slik som studenter, forskere, klinikere og eksterne medarbeidere, vil ha nytte av håndboka. For ansatte som har administrative spørsmål knyttet til oppsummeringsprosjekter, anbefaler vi «Prosjektleders administrative håndbok».

Systematisk oppsummering av forskning kjennetegnes ved en systematisk og eksplisitt framgangsmåte i å formulere spørsmål, søke etter litteratur, vurdere, sammenstille og presentere den forskningsbaserte kunnskapen. Dette er et område i kontinuerlig metodeutvikling der område for helsetjenester i Folkehelseinstituttet både bidrar til- og henter kunnskap internasjonalt. Håndboken synliggjør prinsippene vi jobber etter, men uten å gå helt ned i detaljene på metodene. En grundigere innføring i metodene vi bruker for å utarbeide systematiske oversikter er gitt i håndbøker utgitt av Centre for Reviews and Dissemination (1) og Cochrane (2). Det finnes også andre relevante metodebøker (f.eks. 3-6).

Vi ønsker at alle som arbeider med oppsummering av forskning i Folkehelseinstituttet skal ha et eierforhold til håndboken. Derfor har også svært mange ansatte bidratt til både tidligere utgaver og til denne nye versjonen.

Takk til alle medarbeidere for bidragene.

Kåre Birger Hagen
Fagdirektør

1 Våre ulike oversikter

Vi utarbeider ulike typer systematiske kunnskapsoversikter. Et felles kjennetegn er at de systematisk samler sammen resultater fra eksisterende studier.

Gode kunnskapsoversikter er utarbeidet på en systematisk og transparent måte. Det skal være mulig for andre å etterprøve og kritisere våre metoder, resultater og konklusjoner. Derfor bruker vi eksplisitte og forhåndsdefinerte metoder når vi innhenter, vurderer og sammenfatter forskningsresultater. Grunnlaget for metodene vi benytter følger internasjonale prinsipper for hvordan vi utarbeider systematiske oversikter (1;2). Systematiske oversikter forsøker å svare på et bestemt forskningsspørsmål ved å samle all forskning som passer til forhåndsdefinerte inklusjonskriterier (2).

Systematiske oversikter tar sikte på å redusere systematiske skjevheter ved å benytte eksplisitte og systematiske metoder, og gir dermed mer pålitelige funn som kan bidra til velinformerte beslutninger.

Kunnskapsoversikter kan utføres for mange ulike spørsmål, men vi utfører oftest oversikter for spørsmål om effekt av tiltak, prognose, diagnostikk og erfaring.

PICO er et akronym som hjelper oss å systematisere et forskningsspørsmål som omhandler effekten eller nytten av tiltak. PICO står for Population-Intervention-Comparison-Outcome, og er:

- Populasjon / problem (Hvilke personer handler det om? Hva er problemet?)
- Intervensjon/tiltak eller exposure (Hva gjør vi med dem? Hva blir de utsatt for?)
- Comparison (Hva sammenligner vi med?)
- Outcome/endepunkter (Hvilke/t utfall er av interesse? Hva er det av virkninger man ønsker kunnskap om,- omfatter også negative virkninger).

Akronymet, i helhet eller delvis, kan brukes for også andre spørsmål enn effekt.

Kort fortalt inneholder de fleste av våre oversikter:

- klar tittel og formål
- omfattende referert søkestrategi etter litteratur
- klare kriterier for inklusjon og eksklusjon av studier

- vurdering av risiko for systematiske skjevheter på inkluderte studier
- beskrivelse av karakteristiske trekk for hver av de inkluderte studiene
- sammenstilling av resultatene fra de inkluderte studiene

Område for helsetjenester i Folkehelseinstituttet utfører ulike typer kunnskapsoversikter avhengig av hva forslagsstiller ønsker, hva oversikten skal brukes til og tidsrammen for arbeidet. Tabell 1 gir en kort beskrivelse over våre ulike oversikter og hva de kan brukes til. Resten av kapittelet beskriver oversiktene mer utfyllende.

Tabell 1: Kort beskrivelse over våre oversikter og hva de kan brukes til

Type oversikt	Beskrivelse	Hensikt / bruk
Systematisk oversikt (over primærstudier) (systematic review)	En systematisk oversikt over et klart definert forskningsspørsmål (vanligvis om effekt: er tiltaket virksomt og trygt?)	Gi svar på en problemstilling; bidra til anbefalinger om tiltak; vise hvorvidt/hvor det mangler primærforskning
Systematisk oversikt over oversikter (overview of reviews)	En systematisk oversikt over et bredt forskningsspørsmål der det fins systematiske oversikter som dekker hele eller deler av problemstillingen (vanligvis om effekt: er tiltaket virksomt og trygt?)	Gi oversikt over en problemstilling; bidra til anbefalinger om tiltak; vise hvorvidt det mangler oppsummert forskning
Hurtigoversikt (rapid review)	En systematisk oversikt over et klart definert forskningsspørsmål (vanligvis om effekt: er tiltaket virksomt og trygt?), men med mindre omfattende metode enn ved en systematisk oversikt og dermed mindre pålitelighet. Utføres vanligvis kun når det er svært kort tidsfrist	Som systematisk oversikt (eller oversikt over oversikter), men med mindre pålitelighet pga. mindre omfattende metode
Fullstendig metodevurdering (health technology assessment [HTA])	En systematisk oversikt over et klart definert forskningsspørsmål om effekt med minst ett av følgende tillegg: helseøkonomisk evaluering, vurdering av etiske, juridiske og/eller organisatoriske og sosiale konsekvenser	Gi oversikt over en problemstilling pluss vurdering av økonomiske, etiske, juridiske og/eller organisatoriske og sosiale konsekvenser; bidra til anbefaling/beslutning om tiltak; vise hvorvidt/hvor det mangler forskning
Hurtig metodevurdering (single technology assessment STA)	En systematisk oversikt over et klart definert forskningsspørsmål som vurderer dokumentasjonen for effekt, sikkerhet og kostnadseffektivitet av en ny metode der dokumentasjonen ofte er sammenfattet av produsenten av den nye metoden	Gi oversikt over primærforskningen om en enkelt metode mht. effekt, sikkerhet og kostnadseffektivitet; bidra til anbefaling/beslutning om innføring/finansiering av metoden
Metodevarsel (early warning report / horizon scanning report)	En kort omtale av ny eller etablert metode antatt å være særlig viktig og aktuell for innføring eventuelt utfasing	Gi en kort beskrivelse, gjøre et enkelt systematisk litteratursøk og vurdering av metoden; gi grunnlag for å avgjøre hvorvidt fullstendig metodevurdering bør utføres
Systematisk kartleggingsoversikt (systematic mapping review)	En systematisk oversikt over et klart definert forskningsspørsmål som kartlegger og beskriver hva som fins av litteratur (vanligvis forskningslitteratur) på et spesifikt tema	Gi oversikt over et spesifikt tema; vise hva som fins og ikke fins av forskning; bidra til å avgjøre om det er hensiktsmessig å utføre en systematisk oversikt på et smalere spørsmål innen temaet
Systematisk litteratursøk med sortering (systematic literature search with sorting)	En oversikt over et klart (ofte bredt) definert forskningsspørsmål, men med mindre omfattende metode enn ved en systematisk oversikt, som kun gir en sortert liste over forskningslitteratur	Vise hva som fins og ikke fins av forskning om et definert spørsmål; bidra til å avgjøre om det er hensiktsmessig å utføre en systematisk oversikt
Omtale av systematisk oversikt (presentation of a systematic review)	En kort (1-3 sider) beskrivelse av noen andres systematiske oversikt, med formidling av resultater og GRADE vurdering av tilliten til resultatene. Den kan inneholde tilleggsopplysninger om norske forhold	Gi opplysninger om forventet effekt av tiltak som er relevante for norske forhold

Vi har altså mange ulike produkter i vår portefølje. I tabell 2 beskriver vi de fasene som utføres i utarbeiding av hver av oversiktene.

Tabell 2: Identifisering av de ulike fasene/trinnene som inngår i oversiktene

Trinn	- Systematisk oversikt - Oversikt over oversikter - Fullstendig metodevurdering	Hurtig metodevurdering	Metodevarsel	Systematisk kartleggingsoversikt	Systematisk litteratursøk med sortering	Omtale
Identifisering av litteratur	√	√	√	√	√	√
Utvelgelse	√	√	√	√	√	
Kritisk vurdering	√	vanligvis ikke		vanligvis ikke		
Datauttrekk	√	√		√		
Sortering	√	vanligvis ikke		√	√	
Beskrivelse	√	√	√	√	vanligvis ikke	√
Analyse	√	√		√		
Sammenstilling	√	√				
GRADE	√					√
→Tidsbruk	Ca 12 mnd	Ca 3-6 mnd	Ca 3 mnd	Ca 6 mnd	Ca 3-6 mnd	Ca 1 uke

SYSTEMATISK OVERSIKT

En systematisk oversikt er en litteraturstudie der forfatterne har brukt en systematisk og tydelig framgangsmåte for å finne, velge ut, vurdere og oppsummere all forskning om et definert forskningsspørsmål (vanligvis om effekt av tiltak).

En systematisk oversikt er karakterisert ved at den:

- inneholder et tydelig formål med klare inklusjonskriterier for studier
- har en tydelig, reproducerbar framgangsmåte, inkludert at den bygger på en protokoll som vi utarbeider før vi går i gang med å utføre den systematiske oversikten
- har en oppgitt systematisk søkestrategi som forsøker å finne alle studiene som kan møte inklusjonskriteriene
- vurderer validiteten av de inkluderte studiene, f.eks. ved vurdering av risiko for systematiske skjevheter
- gir en systematisk presentasjon og sammenstilling av kjennetegn og resultater fra de inkluderte studiene

Formålet med en systematisk oversikt er å:

- gi en oversikt over et bestemt forskningsspørsmål
- bidra til å informere anbefalinger eller beslutninger om tiltak

- vise hvorvidt det mangler forskning, og hva slags forskning som det eventuelt er behov for

Vi utfører ikke en systematisk oversikt av primærforskning dersom det allerede foreligger en god og oppdatert systematisk oversikt om den problemstillingen vi er bedt om å undersøke.

Vi velger forskningsdokumentasjon av så høy kvalitet (så lav risiko for systematiske skjevheter) som mulig for problemstillingen. Når vi vurderer spørsmål om effekt av tiltak vil vi derfor i de fleste tilfellene først søke etter randomiserte kontrollerte studier. Dersom slike studier ikke finnes eller problemstillingen krever det, vil vi også søke etter studier med andre typer studiedesign, slik som kontrollerte før- og etter-studier, avbrutte tidsserieanalyser, kasus-kontrollstudier, kohortstudier eller pasientserier.

Vi henter ut data fra primærstudiene som er relevant for problemstillingen vår. Dersom det er mulig og hensiktsmessig, gjør vi samleanalyser (metaanalyser).

Når vi har tid og anledning, forsøker vi å utarbeide systematiske oversikter innenfor rammene av Cochrane og Campbell samarbeidet. Dermed bidrar vi til den internasjonale dugnaden med å oppsummere forskning. Slike oversikter følger de respektive prosessene i Cochrane og Campbell samarbeidet.

OVERSIKT OVER SYSTEMATISKE OVERSIKTER

En oversikt over systematiske oversikter ('paraplyoversikt') er en systematisk oversikt over et bredt forskningsspørsmål der det fins systematiske oversikter som dekker hele eller deler av problemstillingen. Problemstillingen er bredere enn for en systematisk oversikt og omhandler vanligvis effekt der vi spør: er tiltaket virksomt og trygt? En oversikt over oversikter gir svar på hvilke tiltak som har effekt, men som regel ikke hvilke tiltak som har best effekt.

Formålet med en paraplyoversikt er det samme som formålet med en systematisk oversikt, nemlig å:

- gi oversikt over et bestemt forskningsspørsmål
- bidra til å informere anbefalinger og beslutninger om tiltak
- vise hvor det mangler oppsummert forskning, og hvor det mangler primærforskning

Det kan være aktuelt å utføre en paraplyoversikt når søket etter litteratur har vist at det fins flere systematiske oversikter som dekker spørsmålet, der søket ikke er eldre enn 5 år, og oversiktene er av høy eller moderat metodisk kvalitet, slik at vi kan ha tillit til resultatene. (For enkelte spørsmål kan det være relevant

å inkludere også oversikter der søket er eldre enn 5 år, dette bør vurderes i hvert enkelt tilfelle).

Det er mest hensiktsmessig å utføre en paraplyoversikt dersom vi har en problemstilling med:

1. én populasjon og flere tiltak
2. flere populasjoner og ett tiltak
3. én populasjon og ett tiltak men flere utfallsmål
4. flere populasjoner og flere tiltak

HURTIGOVERSIKT

En hurtigoversikt er en systematisk oversikt over et klart definert forskningsspørsmål (vanligvis om effekt og sikkerhet av tiltak), men med mindre omfattende metode enn ved en systematisk oversikt og dermed mindre pålitelighet. En slik oversikt utføres vanligvis kun når det er behov for svært rask leveranse. For å utføre oversikten innen kort tid vil vi f.eks. kunne definere en smalere PICO, ha mindre omfattende søkestrategi, søke i færre litteraturbaser, ha kun én person som gjør utvelgelsen, utelate gradering av vår tillit til dokumentasjonen, eller kun narrativt beskrive de viktigste resultatene.

METODEVURDERING

En metodevurdering (tidligere kalt HTA-rapport og omfatter både fullstendig metodevurdering og hurtig metodevurdering) er resultat av å oppsummere tilgjengelig dokumentasjon om effekt og sikkerhet av en ny eller etablert metode og vurdere konsekvenser av å innføre den nye metoden eller slutte med den gamle ved å vurdere økonomiske, etiske, sosiale, organisatoriske og/eller juridiske aspekter. Ytterligere detaljer om metodevurderinger fins i kapittel 12.

Litteraturgrunnlaget kan være enten systematiske oversikter eller primærstudier. Metodene for systematisk vurdering av effekt og sikkerhet følger samme metode som beskrevet for systematiske oversikter.

Formålet med en metodevurdering er å:

- gi en oversikt over et spørsmål, samt vurdere økonomiske, etiske, juridiske og/eller organisatoriske og sosiale konsekvenser
- bidra til å informere beslutninger om tiltak i Nye metoder
- vise hvor det mangler forskning

Hurtig metodevurdering

En hurtig metodevurdering er en kunnskapsoppsummering med fokus på effekt, sikkerhet og kostnadseffektivitet om ett enkelt legemiddel eller et ikke-

medikamentelt tiltak. Område for helsetjenester i Folkehelseinstituttet utfører hurtig metodevurderinger på metoder som ikke er legemidler.

Ved en hurtig metodevurdering er det produsenten som sender inn dokumentasjon og utarbeider nødvendige analyser. Vi vurderer innsendt dokumentasjon og utarbeider metodevurderingen innenfor en frist på 180 dager. Det utarbeides ikke en prosjektplan.

Helseøkonomisk evaluering

Helseøkonomiske evalueringer er tverrfaglige prosjekter hvor man identifiserer, måler, verdsetter og sammenlikner helsekonsekvenser og kostnader av alternative medisinske tiltak. Hensikten er å legge til rette for en størst mulig helseeffekt av de ressursene som helsevesenet har til rådighet. Vi utarbeider slike evalueringer som del av metodevurderinger. Evalueringene er i de fleste tilfeller basert på egenutviklede modeller for sykdomsforløp og behandling.

Helseøkonomisk oversikt

Helseøkonomiske oversikter er resultat av systematisk søk etter og oppsummering av helseøkonomiske evalueringer utført av andre. En helseøkonomisk oversikt er del av en metodevurdering. En helseøkonomisk oversikt blir laget når det er hensiktsmessig, eller nå vi ikke har kapasitet til å lage en helseøkonomisk evaluering (se avsnittet over).

METODEVARSEL

Metodevarsler er korte omtaler av nye metoder antatt å være særlig viktige og aktuelle for innføring i Norge. Oppgaven består i å identifisere, beskrive og følge nye metoder. Aktiviteten omfatter en filtreringsprosess basert på om metoden er ny, relevant og viktig, og en undersøkelse av om det foreligger klinisk forskningsdokumentasjon som kan gi grunnlag for en metodevurdering. Metodevarslene publiseres på www.helsebiblioteket.no/mednytt.

Formålet med et metodevarsel er å:

- gi en kort beskrivelse og vurdering av en ny metode
- gi Bestillerforum grunnlag for å avgjøre om det er nødvendig å utarbeide en metodevurdering

SYSTEMATISK KARTLEGGINGSOVERSIKT

En systematisk kartleggingsoversikt er en systematisk oversikt over et klart definert forskningsspørsmål som kartlegger og beskriver hva som fins av litteratur (vanligvis forskningslitteratur) på et spesifikt tema (Arksey and O'Malley 2005; Levac, Colquhoun, og O'Brien, 2010).

Slike oversikter gir et overblikk over litteraturen som fins på et fagområde, slik at vi bedre kan vurdere hvordan vi skal utvikle feltet videre. Forskerne benytter systematiske og eksplisitte metoder for å identifisere, utvelge, og beskrive relevant litteratur/forskning om et tema.

Trinnene er: 1) identifisere problemstillingen, 2) identifisere relevante studier, 3) velge studier, 4) kartlegge data, 5) sette sammen, oppsummer, rapportere. Et siste trinn, konsultasjon, er valgfritt. Et eksempel på et spørsmål for en systematisk kartleggingsoversikt er «Hva fins av og kjennetegner forskning om tvangsbehandling i psykisk helsevern?»

Formålet med en systematisk kartleggingsoversikt er å:

- summere og beskrive litteraturflommen med hensyn til et spesifikt tema (f.eks. hva slags studier som er utført, populasjoner som er studert, utfall som er målt osv.)
- klargjøre hva som fins og ikke fins av forskning (avdekke kunnskapshull og vise vei for videre forskning)
- bidra til å avgjøre om det er hensiktsmessig å utføre en systematisk forskningsoversikt på (deler av) spørsmålet

SYSTEMATISK LITTERATURSØK MED SORTERING

Et systematisk litteratursøk med sortering er en oversikt over et klart definert forskningsspørsmål, som gir en sortert liste over forskningslitteratur. Metoden ved systematisk litteratursøk med sortering er mindre omfattende enn for de andre oversiktene. Forskningslitteraturen blir vanligvis ikke lest i fulltekst, kun titler og sammendrag. Vi gjennomfører heller ingen kvalitetsvurdering, beskrivelse av eller sammenstilling av litteraturen. Vi sorterer mulig relevant litteratur i grupper for å vise hvilken forskning som fins om spørsmålet.

Litteratursøk med sortering kan være et utgangspunkt for en mer presis bestilling av en systematisk oversikt eller en metodevurdering.

Formålet med et systematisk litteratursøk med sortering er å:

- gi en sortert liste over forskningslitteraturen om et forskningsspørsmål
- vise hva som fins og ikke fins av forskning om spørsmålet
- vise hvorvidt/hvor det mangler forskning, og eventuelt hva slags forskning som mangler
- bidra til å avgjøre om det er hensiktsmessig å utføre en systematisk forskningsoversikt om spørsmålet

OMTALE AV SYSTEMATISK OVERSIKT

En omtale av en systematisk oversikt er en kort (1-3 sider) beskrivelse og GRADE vurdering av systematiske oversikter utarbeidet av andre. Den kan inneholde tilleggsopplysninger om norske forhold.

Omtaler presenterer gjerne systematiske oversikter utgitt av Cochrane eller Campbell samarbeidet, men også relevante metodevurderinger eller oversikter som er publisert av andre.

Omtaler er nyttige fordi Folkehelseinstituttet kun kan besvare en liten del av alle spørsmål som eiere, utøvere og brukere av norsk helsetjeneste ønsker svar på ved hjelp av oppsummert forskning. I mange tilfeller er det derfor god ressursutnyttelse å formidle en allerede eksisterende oversikt fra en av våre internasjonale samarbeidspartnere.

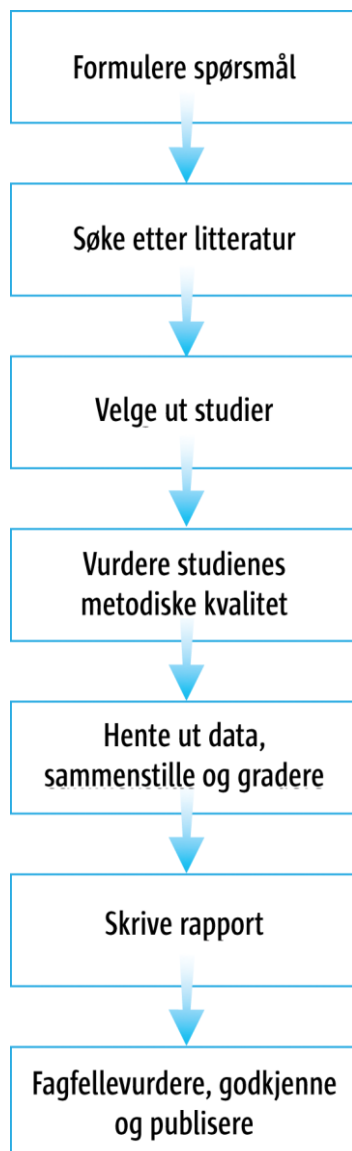
Omtalene blir gjennomgått av norske fagfolk på området før publisering. Vi publiserer våre omtaler på fhi.no, via våre nyhetsbrev og i andre relevante kanaler, blant annet Tidsskrift for jordmødre, Sykepleien Forskning og Helsebiblioteket.

Formålet med en omtale er å:

- gi en kort beskrivelse og vurdering av en oversikt
- gi opplysninger om tiltak som er relevante for norske forhold

2 Overblikk: Slik oppsummerer vi forskning

Når vi oppsummerer forskning, går vi gjennom flere faser uansett hvilket produkt / hva slags oversikt vi skal ende opp med. Innholdet i de ulike fasene varierer litt for arbeidet med ulike typer oversikter (se tabell 2 over). Figuren nedenfor viser fasene skjematisk.



Figur 1: Fasene i oppsummering av forskning

I kapittel 3-10 presenterer vi detaljer for disse fasene slik de gjennomføres ved en systematisk oversikt om effekt av tiltak. Deretter tar kapittel 11 for seg diagnostikk, kapittel 12 beskriver metodevurderinger og kapittel 13 beskriver gjennomføringen av en oversikt med kvalitativ forskning.

HVEM GJØR HVA I PROSESSEN?

Praktisk talt alle oversikter lages av en gruppe personer; ofte i et samarbeid mellom eksterne fagpersoner og medarbeidere ved Folkehelseinstituttet. Vi ønsker også å få innspill fra brukere (helsepersonell, pasienter og pårørende), slik at oversikten er relevant for dem som spørsmålet angår mest.

Alle prosjekter har en (intern) prosjektleder som er ansvarlig for prosessen. En leder i avdelingen er ansvarlige for faglig støtte og veiledning, enten ved at han/hun bistår selv eller ved å finne andre som kan ta denne rollen. Minst én annen forsker bidrar i arbeidet med å utføre oversikten.

Folkehelseinstituttets bibliotekarer spiller en sentral rolle, i første rekke ved utarbeiding av søkestrategi (i samarbeid med prosjektgruppa) og deretter med søk etter litteratur. Statistikere bidrar når det er aktuelt med en kvantitativ sammenstilling av resultater (metaanalyse).

Dersom vi skal gjøre en økonomisk analyse, vil en helseøkonom være med og vurdere omfang og valg av type analyse. Helseøkonomen vil også stå for selve analysen, enten alene eller i samarbeid med en utredningsgruppe eller andre eksterne fagfolk.

Dersom vi skal gjøre etiske, juridiske eller organisatoriske vurderinger, bør vi ta kontakt med og få innspill fra fagpersoner på de aktuelle feltene.

Bruk av eksterne fagpersoner og brukere

I hvilken grad vi involverer eksterne medarbeidere/eksperter og brukere/pasientrepresentanter ved utarbeiding av oversikter, varierer mye, og det blir vurdert fra prosjekt til prosjekt. FHI anser brukerinvolvering og deltakelse som en styrke for alle helseforskningsprosjekter.

Noen ganger er det hensiktsmessig å etablere en gruppe som deltar i hele prosessen. Gruppen kan bidra til prosjektplan, ved seleksjon av studier, vurdering av metodisk kvalitet og i skrivning. En slik gruppe består gjerne av eksterne fagpersoner med særskilt innsikt i temaet, brukerrepresentanter og av personer fra område for helsetjenester i Folkehelseinstituttet. Vi forsøker å invitere eksterne fagpersoner med en viss faglig og geografisk spredning for å gjenspeile mangfold eller uenighet som det kan være viktig å ta hensyn til i

prosjektet. Det er viktig at deltakere som inviteres til å delta i slike grupper forplikter seg til å levere sine bidrag til avtalte frister. Andre ganger vil vi knapt involvere eksterne fagpersoner, for eksempel når det er snakk om korte tidsfrister.

En mellomløsning er å etablere en utredningsgruppe som deltar i deler av prosessen. En slik gruppe kan for eksempel bidra i startfasen når vi skal formulere spørsmål, ved utforming av søkestrategi og ved diskusjon av resultater. Vi gjør ofte seleksjon, vurdering og sammenstilling av studier internt i Folkehelseinstituttet.

Systematiske oversikter fra Cochrane og Campbell samarbeidet

Folkehelseinstituttets medarbeidere deltar ofte i utarbeidelsen av Campbell- og Cochrane-oversikter. Prosjektgruppa som utarbeider en Cochrane- eller Campbell-oversikt består som oftest av fagfolk fra flere land med både tematisk og metodisk spisskompetanse. Man prøver også å få med brukere i gruppene, alternativt blir brukere invitert til å gi innspill til utkast underveis.

Folkehelseinstituttets gruppe koordinerer sitt arbeid mot de aktuelle oversiktsgruppene (review groups) i Cochrane- og Campbell-samarbeidet. Det finnes systemer for registrering av titler, godkjenning av protokoller og ferdige oversikter, og forfatterne vil få hjelp underveis. Både protokoll og utkast til den endelige systematiske oversikten gjennomgår ekstern fagfelleevaluering.

3 Formulere spørsmål og skrive prosjektplan



FORMULERE SPØRSMÅL

Det er viktig innledningsvis å avklare hva oppdragsgiver egentlig vil ha svar på. Som en del av utvikling av prosjektplanen vurderer vi i samråd med forslagsstiller derfor hvilke spørsmål som det kan være aktuelt å søke svar på. Dersom vi jobber sammen med en ekstern eller intern arbeidsgruppe, drøfter vi både avgrensning av tema(er) og strukturering (av spørsmål) med den. Oppdragsgiver/forslagsstiller bør alltid få mulighet til å respondere på våre forslag til omformulering og presisering av det opprinnelige forskningsforslaget. Det er viktig med et tett samarbeid med forslagsstiller i startfasen for å få en felles forståelse av hva prosjektet skal gi svar på, og hva svarene i oversikten skal brukes til.

Vi ønsker å være tydelige på hvilke deler av et tema vi tar for oss, f.eks. at en oppsummering handler om diagnostikk, forebygging og behandling, men ikke om årsaksforhold eller organisering av tjenester. I et annet tilfelle kan nettopp spørsmålet være hva den beste måten å organisere en tjeneste på er, mens vi lar spørsmål om effekt av ulike former for behandling ligge.

For å skape klarhet i hvilke spørsmål vi skal besvare slik at vi kan bygge en søkestrategi og kombinere søkeord riktig, er det viktig å strukturere, eller dele opp, spørsmålet eller spørsmålene. En vanlig måte å systematisere et forskningsspørsmål som omhandler effekten eller nytten av tiltak er å formulere spørsmålet i en såkalt PICO. Akronymet PICO står for Population-Intervention-Comparison-Outcome, og er:

- Populasjon / problem (Hvilke personer handler det om? Hva er problemet?)
- Intervensjon/tiltak eller exposure (Hva gjør vi med dem? Hva blir de utsatt for?)
- Comparison (Hva sammenligner vi med?)
- Outcome/endepunkter (Hvilke/t utfall er av interesse? Hva er det av virkninger man ønsker kunnskap om,- omfatter også negative virkninger).

For å definere PICO kan det hjelpe å fylle ut en «tabell» som vist under. Her står P for populasjon (population), I for tiltak (intervention), C for sammenlikning (comparison) og O for utfall/endepunkt (outcome).

Er trening et effektivt tiltak for å redusere overvekt hos barn sammenlignet med råd om kosthold?

Spørsmål om EFFEKT	Hvem (P)	Tiltak (I)	Alternativt tiltak (C)	Utfall (O)
	Overvektige barn	Trening	Råd om kosthold	Vektutvikling

Er perkutan koronar intervensjon (PCI) mer kostnadseffektivt enn trombolyse ved hjerteinfarkt?

Spørsmål om KOSTNADS-EFFEKTIVITET	Hvem (P)	Tiltak (I)	Alternativt tiltak (C)	Utfall (O)
	Pasienter med førstegangsinfarkt	Perkutan koronar intervensjon (PCI)	Trombolyse	Kostnads-effektivitet

Er computertomografi (CT) med kontrast bedre enn urografi for å vurdere pasienter med mistenkt nyrestein?

Spørsmål om DIAGNOSE	Hvem (P)	Diagnostisk test (I)	Ev. alternativt tiltak (C)	Utfall (O)
	Pasienter med mistenkt nyrestein	CT med kontrast	Urografi	Testens nøyaktighet

Har barn som er født for tidlig, økt risiko for å utvikle lærevansker sammenlignet med barn som er født ved termin?

Spørsmål om ÅRSAK	Hvem (P)	Risikofaktor (I)	Sammenlikning (C)	Utfall (O)
	Barn i skolealder	Født for tidlig	Født ved termin	Lærevansker

Hvordan opplever barn og ungdom som bor i institusjon kontinuitetsfremmende tiltak i institusjon?

Spørsmål om OPPLEVELSE	Hvem / Perspektiv (P)	Interesse (I)	Kontekst (K)
	Barn og ungdom som bor i institusjon	Kontinuitetsfremmende tiltak	Vestlige land

Å strukturere spørsmålene på denne måten tydeliggjør hva en rapport faktisk skal handle om, og det tvinger forfatterne vekk fra overordnede temabeskrivelser til konkrete spørsmål. Før vi kan utforme en konkret søkestrategi, må vi imidlertid gjøre begrepene tydeligere. For eksempel må vi gi begrepene «født for tidlig» og «overvektige barn» en form som vi kan behandle vitenskapelig (operasjonalisere) før vi kan bruke uttrykkene til å søke etter litteratur.

I denne fasen er det også viktig å tydeliggjøre hvilke sammenlikninger vi skal gjøre. For eksempel bør vi presisere om trening bare skal sammenliknes med råd om kosthold eller om vi også skal sammenlikne trening med ingen tiltak, eventuelt kombinasjonen av trening og kostholdsrad versus ingen tiltak. Vi må

også beskrive hvilke utfall vi er mest opptatte av. En slik sortering medfører at én og samme studie kan bidra i flere sammenlikninger i den endelige oversikten.

Spørsmålet avgjør hva slags forskning vi leter etter

Vi skiller mellom ulike typer spørsmål. Vi kaller dem gjerne helsetjenestens kjernespørsmål (3;4;6-8). Hvilke former for forskningsbasert kunnskap eller framgangsmåter i forskningen, også kalt studiedesign, som gir gyldig kunnskap om et fenomen, avhenger av spørsmålet. Hvis spørsmålet handler om effekt av forebygging, er det foretrukne design et annet enn om vi er opptatt av hvor treffsikker en ny diagnostisk metode er eller hvordan brukere opplever et spesielt tiltak. Tabell 3 viser hva som er foretrukket design for de ulike kjernespørsmålene.

Tabell 3: Helsefaglige kjernespørsmål og måter man går fram på i forskning for å svare på dem

Kjernespørsmål	Foretrukket design
Hvor mange har en sykdom eller et problem? (spørsmål om forekomst)	Tverrsnittstudie
Hvorfor blir noen syke, mens andre holder seg friske? (spørsmål om årsak)	Kohortstudier eller kasus-kontroll-design
Hvordan avgjøre om en person er syk eller har et bestemt problem? (spørsmål om diagnostikk)	Tverrsnittsundersøkelser (med en referansestandard)
Hva er positive og negative effekter av forebygging, behandling, organisering av tjenester eller andre tiltak? (spørsmål om effekt)	Randomiserte kontrollerte forsøk
Hvordan vil det sannsynligvis gå med en pasient med en bestemt lidelse? (spørsmål om prognose)	Kohortstudier
Hvordan oppleves det å ha dette problemet / være i denne situasjonen? (spørsmål om erfaringer, opplevelser)	Kvalitative studier

Dette er «kjøreregler» som gjelder for forskere som skal planlegge en studie, og som vi følger når vi oppsummerer forskning. En slik oppstilling er imidlertid ikke fullt ut dekkende. For eksempel er det for årsaksspørsmål også slik at et randomisert kontrollert forsøk ofte hadde gitt mer pålitelige sammenlikninger (f.eks. mellom røykere og ikke-røykere med hensyn til kreftutvikling). Men et eksperimentelt design er ofte uetisk når det man undersøker er sammenhengen mellom en mistenkt påvirkbar risikofaktor og utvikling av sykdom. Også når det gjelder spørsmål om effekt av virkemidler, kan det være uetisk eller praktisk vanskelig å gjennomføre forsøk. Da er det aktuelt å bruke andre design i forskningen og i de studiene vi inkluderer i oppsummering av forskning. Andre design enn dem som er nevnt over kan også være aktuelle, som f.eks. avbrutte

tidsserieanalyser, kontrollerte før-og-etter studier og såkalte regression discontinuity designs. Det finnes flere slike ikke-randomiserte design (9).

Det kan oppstå misforståelser om synet på hva som er pålitelig empiri når det gjelder effekten av tiltak. I prinsippet er det opplagt slik at all erfaring med virkemiddelet (fra den første pasient, klient eller bruker som mottar tiltaket) er noe man kan lære av. Prøver man tiltaket på flere, lærer man mer. Man sammenlikner f.eks. med «hvordan det gikk med liknende pasienter før». Har forskeren en parallell gruppe som får et annet tiltak, lærer man enda mer. Eller mer presist: Man prøver på ulike måter å beskytte seg stadig bedre mot å ta feil når man fastslår at det er en kausal sammenheng mellom bestemte tiltak og bestemte virkninger. Ofte er det så mye annet som kan ha skapt virkningen, at det er behov for å gjøre forskningen på måter som best sikrer at man trekker slutninger på basis av «rettferdige sammenlikninger». Det å fordele personer tilfeldig til å få et tiltak mens andre ikke får det, er en rask og effektiv måte å lage en rettferdig sammenlikning på. Ulike kjennetegn ved personene (kjente og ukjente faktorer som vil virke inn på utfallet) vil da være likt fordelt i de to gruppene, og det er mulig i ettertid å trekke slutninger om at det faktisk var tiltaket som ga en eventuell effekt.

Av og til trenger man ikke forskning for å komme til klarhet i hvordan ting henger sammen. Det gjelder særlig når tiltaket har umiddelbar og dramatisk effekt i en situasjon der det tidligere viste seg alltid å gå på en bestemt måte. Et enkelt eksempel er tilførsel av insulin ved altfor høyt sukkerinnhold i blodet. Da er forskning overflødig; det er nok å lære av den umiddelbare erfaringen. Gang på gang i vitenskapshistorien har det imidlertid vist seg at faren for å ta feil når man forsøker å avgjøre om «noe» har ført til «noe annet», generelt er stor. Systematisk innhentet erfaring (dvs. forskning) er derfor som regel bedre enn usystematisk praksiserfaring for å bedømme sammenhenger, og forskning der andre forskjeller (enn at noen mottok et tiltak) er jevnet ut, er bedre enn forskning der sammenlikningsgrunnlaget er svakere.

Kvalitativ forskning har en viktig plass i helsefaglig og samfunnsvitenskapelig forskning. Det gjelder når målet er å skape innsikt i et fenomen der det ikke finnes noen «sannhet»; f.eks. hvordan foreldre opplever å miste et nyfødt barn. Og det gjelder når man forsøker å forklare «hvorfør det gikk som det gikk» i andre typer forskning; f.eks. når man lurer på hvorfor et tiltak viste seg ikke å ha effekt. Kvalitative og kvantitative design kan derfor ofte supplere hverandre. Et eksempel kan være når man bruker dybdeintervjuer for å karakterisere noen former for reaksjoner på å miste et barn, spørreundersøkelser for å finne ut om disse reaksjonene er utbredte, fokusgrupper for å komme på sporet av mulige virkemidler som kan gjøre det lettere for foreldrene å mestre sorgen, kontrollerte forsøk for å vurdere effekten av f.eks. sorggrupper og igjen kvalitative design for å forsøke å forstå hvorfor slike tiltak har eller ikke har effekt (eller hvorfor de virker for noen og ikke for andre).

Samlet sett er det grunn til å utvise ydmykhet når vi forsøker å forstå hva som «er i verden», for det er lett å ta feil. Forskning og ulike forskningsdesign, inkludert det å lage systematiske oversikter, er forsøk på sette denne ydmykheten i system og minske faren for at man tar feil.

Endelig er det grunn til å understreke at innsikt i et bestemt felt, f.eks. gjennom lang praksiserfaring, er viktig for å tolke forskningsresultater. Det kan f.eks. hende at forskningen er gjennomført på så atypiske pasienter at man egentlig ikke har lært noe om situasjoner som oppstår i «det virkelige liv». I tillegg til å bedømme forskningens gyldighet er det altså viktig også å bedømme læringsverdien eller overførbarheten.

Seleksjonskriterier

I seleksjonen av studier er det viktig å unngå systematiske feil. Derfor er det viktig at avgjørelsene om inklusjon og eksklusjon er basert på forhåndsbestemte kriterier som er spesifisert i prosjektplanen.

Inklusjons- og eksklusjonskriterier skal logisk følge av spørsmålet oversikten skal besvare. Vi definerer dem ved å være tydelige på hvilke personer, intervensjoner (tiltak), sammenlikninger, utfall og studiedesign vi ønsker å dekke i oversikten (og hvilke vi ikke er interesserte i).

Det er ofte diskusjon om hvor brede eller smale disse kriteriene skal være. Tar vi med få studier, reduseres ofte nytten av oppsummeringen. Tar vi med mange ulike typer studier, får vi ofte med informasjon det er vanskelig å sammenlikne og oppsummere.

Vi har med studiedesign som kriterium fordi vi ønsker å basere oppsummeringen på «den beste kunnskapen». Når vi arbeider på et område der det ikke foreligger mange, godt utførte studier med adekvat design som har undersøkt spørsmålet oversikten skal svare på, må vi avveie fordeler og ulemper ved å ta med studiedesign der vi vet at faren for feilslutninger er større.

Tabell 4: Eksempel på seleksjonskriterier

	Inklusjonskriterier	Eksklusjonskriterier
Spørsmålet: Vil antibiotikaprofylakse redusere sjansen for postoperative infeksjoner blant pasienter som hofteopereres?		
Populasjon	Pasienter som hofteopereres (primært eller ved reoperasjon)	Andre operasjoner
Intervensjoner	Antibiotikaprofylakse	Annen antibiotikabruk enn profylakse
Sammenlikning	Placebo eller ingen profylakse	

Utfall	Infeksjon i operasjonssåret bekreftet med adekvate mikrobiologiske teknikker	Infeksjon som ikke er bekreftet
Studiedesign	Randomiserte kontrollerte forsøk	Ikke-randomiserte studier

Restriksjoner på grunnlag av språk

Å begrense inklusjon av studier på grunn av språket publikasjonen er skrevet på, kan introdusere systematiske feil. Det er imidlertid ofte begrenset kapasitet og kompetanse til å oversette artikler, rapporter osv. Det er en hjelp at bl.a. MEDLINE indekserer artikler på norsk, tysk og en del andre språk og at de publiserer engelskspråklig sammendrag. De automatiske oversettelsesprogrammene på nettet, slik som Google oversetter, er også etter hvert blitt så gode at de kan være til hjelp slik at vi kan vurdere ikke-engelske sammendrag for potensiell relevans.

SKRIVE PROSJEKTPLAN

Etter at spørsmålet er godt strukturert må vi utrede spørsmålet og utarbeide en prosjektplan. Dette innebærer først å undersøke hvorvidt det fins andre oversikter om spørsmålet og vurdere hvilken type oversikt som vi skal utarbeide.

Fins det andre oversikter om spørsmålet?

Systematisk oversikter om effekt av tiltak identifiserer, vurderer og oppsummerer resultater fra store mengder forskning. Dette er arbeidskrevende prosesser som pågår verden over. Flere av våre søsterinstitusjoner i andre land og internasjonale nettverk som Cochrane, INAHTA (International Network of Agencies for Health Technology Assessment) og Campbell Collaboration arbeider kontinuerlig for å skape oversikt over eksisterende og ny kunnskap. Det er viktig å unngå dobbeltarbeid, og derfor undersøker vi om det finnes gode systematiske oversikter, før vi går i gang med en ny. Dersom det fins gode oversikter som trenger oppdatering, er det fornuftig å bygge videre på arbeid andre har gjort, i stedet for å starte prosessen fra grunnen av.

Alle systematiske oversikter som påbegynnes, bør meldes inn i Prospero-databasen (se www.crd.york.ac.uk/prospero/) som inneholder protokoller over systematiske oversikter, og i POP-databasen (EUnetHTA). Dette beskytter mot unødig dobbeltarbeid.

I starten av et prosjekt gjennomføres det søk for å identifisere eventuelle systematiske oversikter på området. Det kan også være aktuelt å gjøre testsøk for å bedømme sannsynlig volum av primærstudier. Slike orienterende øvelser, eller temautredninger, (på engelsk «scoping») vil gi erfaring som er nyttig for å bygge opp en søkestrategi.

Hva slags oversikt skal utarbeides?

Folkehelseinstituttet lager flere typer oversikter (jfr. kapittel 2). Av og til er formatet gitt, som når oppdragsgiver trenger en oversikt i løpet av få uker. For øvrig påvirkes formatet av hva som finnes av oppsummert kunnskap fra før og hva oversikten skal brukes til. Dersom andre nylig har laget oversikter på et bestemt område for spørsmål vi er opptatt av å utrede, bygger vi på dem eller formidler dem. Dersom det ikke finnes systematiske oversikter for det aktuelle spørsmålet er det aktuelt for oss å igangsette og registrere en systematisk oversikt eller metodevurdering i de internasjonale nettverkene. Hvilken type oversikt som skal utarbeides avgjøres i dialog med forslagsstiller/oppdragsgiver.

Prosjektplan

Etter at problemstillingen er godt utredet og formulert lages en prosjektplan. En prosjektplan beskriver hva oppsummeringen skal omhandle og hvordan den skal utføres. Planen beskriver relevant bakgrunnsinformasjon, spørsmålet (eventuelt spørsmålene) som skal utredes og framgangsmåten (metoden) for oppsummeringen, samt de viktigste aktivitetene og ressursene som kreves for omfanget som prosjektet har. Prosjektplanen utvikles i dialog med forslagsstiller/oppdragsgiver slik at både forskerne og oppdragsgiverne blir fornøyde med presiseringen av oppdraget. Planen blir brukt av prosjektleder som en baseline som prosjektfremdriften skal måles opp mot. Hva slags prosjektplan som lages, avhenger av hva slags type oversikt som planlegges (jfr. kapittel 2). Omfanget av problemstillingen og tidsrammen til rådighet er avgjørende faktorer for hvordan arbeidet legges opp, i tillegg til hva som eventuelt allerede finnes av oppsummert forskning for den aktuelle problemstillingen. Når vi benytter en ekstern utredningsgruppe, er gruppa medansvarlig for utformingen av prosjektplanen.

Prosjektleders administrative håndbok, som finnes på Folkehelseinstituttets intranett, omhandler tekniske og praktiske aspekter (for eksempel mal for prosjektplan og hvilke dokumenter som skal registreres i arkivet). Dette inkluderer fagfelle vurdering. Fagfelle vurdering av både prosjektplan og den ferdige oversikten er viktig for å kvalitetssikre arbeidet. Prosjektleder bør tidlig skjele til rapportmalen, som beskriver hvordan en rapport skal settes opp og hva den skal inneholde. Malen ligger på Folkehelseinstituttets intranett.

Hva inneholder en prosjektplan?

Bakgrunn

Spørsmål oversikten skal besvare

Søkestrategi; søketermer, ressurser det skal søkes i og angitt tidsrom

Inklusjonskriterier for studier (eventuelt oversikter)

Hvordan studiene skal kvalitetsvurderes

Hvordan data skal trekkes ut og hvilke data som skal trekkes ut

Hvordan informasjonen skal oppsummeres (inklusive analyseplan)

Tidsplanlegging

Godkjenne og publisere prosjektplan

Alle prosjektplaner gjennomgår fagfellevurdering. Hvem og hvor mange fagfeller som vurderer hver prosjektplan avhenger av hvilken type oversikt som utføres.

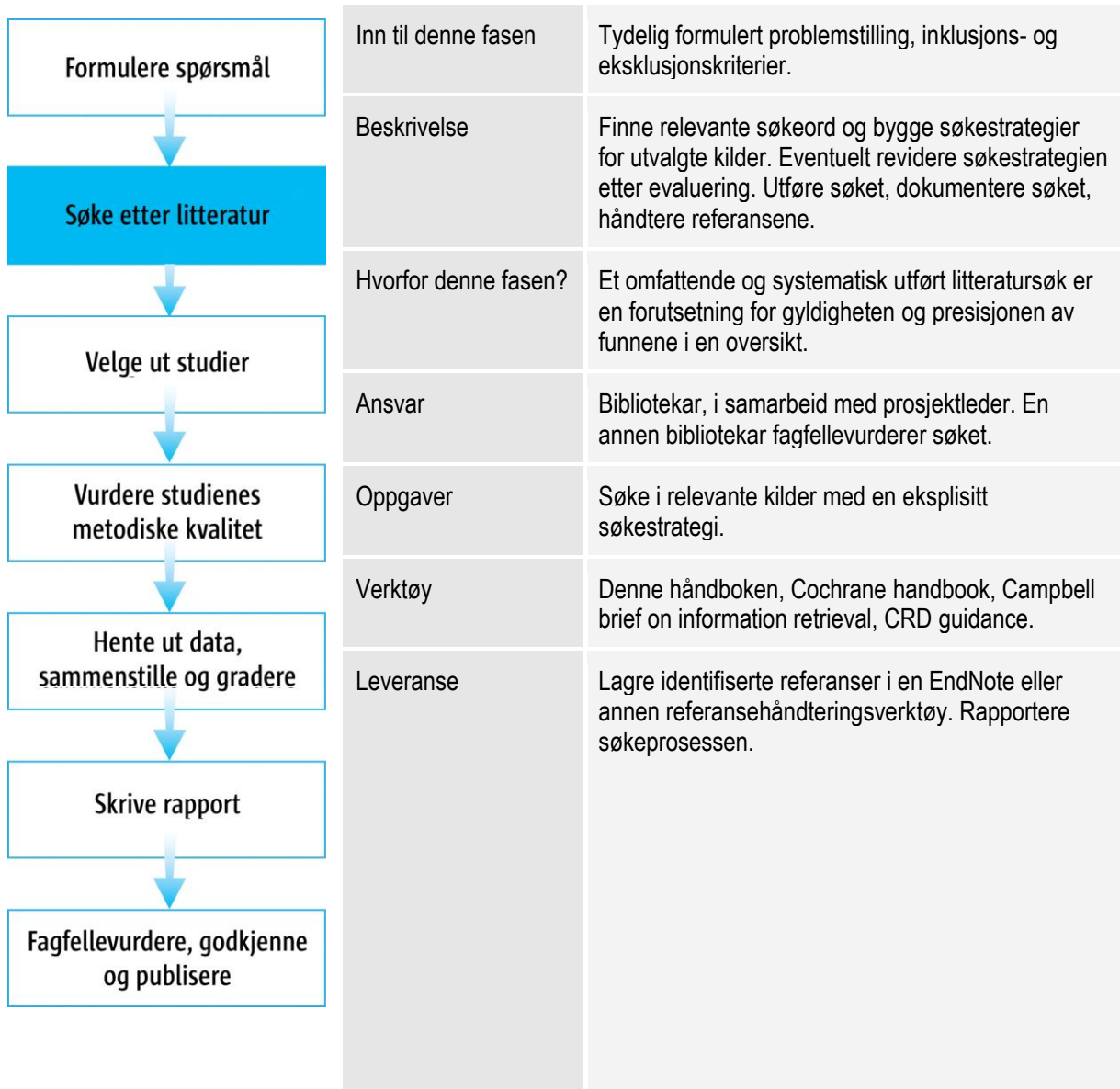
Prosjektplanene blir deretter lagt ut på fhi.no, og det lages i noen tilfeller en kort nettomtale av prosjektene, som også sendes ut med nyhetsbrevet.

Det kan være aktuelt å åpne for kommentarer til prosjektplanen i en avgrenset tid for å få sjekket om utfallsmålene dekker brukernes behov. I så fall skal det nevnes i prosjektplanen.

Viktige trinn når vi utarbeider en prosjektplan

- Å lage en prosjektplan er et avgjørende første skritt i å utføre en oppsummering. En gjennomarbeidet plan, basert på gode metoder for syntetisering av forskningsbasert informasjon, er i seg selv viktig for å sikre god kvalitet i arbeidet. God dialog og samarbeid med oppdragsgiver er i tillegg viktig for å være sikker på at presiseringen av oppdraget er i henhold til det oppdragsgiver ber om.
- Å spesifisere problemet(ne) i form av presist formulerte spørsmål er avgjørende for et godt resultat. Spørsmålene skal formuleres på forhånd (a priori) og beskrive hvilke pasienter (eventuelt befolkningsgrupper), intervensjoner, sammenlikninger, utfall og studiedesign som er aktuelle.
- På basis av denne inndelingen utvikles inklusjons- og eksklusjonskriterier og søkestrategien.
- Dersom spørsmålet handler om å identifisere effekter av et tiltak, er det først og fremst aktuelt å inkludere godt kontrollerte (randomiserte) forsøk eller systematiske oversikter som oppsummerer slike eksperimentelle data.
- Dersom oversikten også skal beskrive prosesser for implementering av et tiltak (for eksempel om en aktuell screeningmetode kan tas i bruk i Norge) og/eller pasienterfaringer, bør vi overveie å inkludere kvalitative studier.
- Hvis oppsummeringen også skal omfatte kostnadseffektvurderinger, inkluderer vi helseøkonomiske evalueringer som både beskriver medisinske og økonomiske effekter.
- Søkestrategien utvikles av en bibliotekar med særskilt innsikt i relevante kilder og hvordan man bygger søkestrategier for hver av dem, i samarbeid med prosjektgruppen.
- Prosjektplanen skal beskrive hvordan utvelgelsen av artikler skal foregå, hvordan man vil kvalitetsvurdere studiene og hvordan data skal hentes ut.
- Prosjektplanen skal beskrive hvordan eventuell sammenstilling og analyse av resultatene skal utføres, og plan for gradering.
- Utkast til prosjektplan bør fagfellevalueres.
- En godkjent prosjektplan bør bare endres dersom det er gode grunner for å gjøre det. Begrunnelsen for endringene bør være med i den endelige publikasjonen slik at endringene kan vurderes av andre. Det kan for eksempel komme inn godt begrunnede forslag fra brukerne om flere eller andre utfallsmål.

4 Søke etter litteratur



HVA KJENNETEGNER GODE LITTERATURSØK?

Et godt utført litteratursøk er viktig for å identifisere all relevant forskning som har vært utført i verden (eller en del av verden) om et bestemt spørsmål.

Grundigheten og framgangsmåten i litteratursøket er et av de trekkene som skiller systematiske oversikter fra tradisjonelle oversiktsartikler. Et omfattende søk etter litteratur styrker oversiktens validitet (gyldighet). Det er viktig at identifiseringen av forskning er så fri for systematiske feil som mulig. For eksempel er noen forskningsområder preget av ulike former for selektiv publisering. Et godt utført litteratursøk identifiserer systematiske oversikter og primærstudier som samsvarer med inklusjonskriteriene, og ekskluderer irrelevant litteratur.

Tabell 5: Trinn i et systematisk litteratursøk

Trinn	Handling	Beskrivelse
1.	Strukturere spørsmålet	Dele spørsmålet inn i kategorier og bestemme hva det handler om.
2.	Velge databaser/kilder	Valget baserer seg på tema, type spørsmål, produkt og i noen tilfeller tidsfrist.
3.	Lage søkestrategier for utvalgte kilder	Finne relevante søkeord for de ulike kategoriene i spørsmålet. Lage søkestrategier for hver av de inkluderte kildene og legge til relevant metodefilter. Gjøre søkene.
4.	Vurdere resultatene og eventuelt revidere søkestrategiene	Hvis søket gir for mange irrelevante treff og/eller utelater relevante referanser, bør søkestrategiene revideres. Nye kilder må eventuelt også vurderes.
5.	Bearbeide referanser	Importere referanser fra ulike kilder til EndNote (eller annet tilgjengelig referanse-håndteringsverktøy), fjerne dubletter m.m.
6.	Loggføre og rapportere søket	Valg som tas underveis i søkeprosessen bør loggføres. Søket beskrives i rapportens metodekapittel, inkludert dato. Søkestrategier legges i vedlegg.
7.	Eventuelt oppdatere søket	Dersom oppsummeringen skal oppdateres eller prosjektet tar lang tid, må også søket foretas på nytt.

Effektiv søking krever særskilt kompetanse, og det er ønskelig å involvere spesialtrente bibliotekarer tidlig. Bibliotekar, prosjektleder og gjerne eksperter på fagområdet bør arbeide sammen for å utvikle en søkestrategi. Bibliotekaren har ansvar for selve søket.

Et godt designet og utført systematisk søk er grunnsteinen i en kunnskapsoppsummering. Derfor har vi i Folkehelseinstituttet som en regel at søkene vurderes av en annen spesialtrent bibliotekar.

VELGE UT DATABASER OG ANDRE LITTERATURKILDER

Hvilke kilder søker vi i?

Hvilke kilder vi søker i avhenger av flere faktorer: type oversikt, type spørsmål (effekt, diagnose osv.), tema og i noen tilfeller tidsfrist. Vi leter vanligvis først i

bibliografiske databaser etter systematiske oversikter og utvider søket til primærstudier dersom oversikter mangler, er utdaterte eller litteratursøket i oversikten(e) ikke er rapportert eller er mangelfullt rapportert.

Spørsmål, studiedesign og valg av kilder henger nøye sammen. Tabell 6 gir en generell oversikt over kjernesporsmål med foretrukket studiedesign og noen foretrukne kilder til systematiske oversikter og primærstudier.

Tabell 6: Kjernesporsmål med foretrukket studiedesign og kilder til litteratur

Spørsmål		Foretrukket studiedesign	Databaser/kilder
Hvor mange...?	Forekomst (prevalens)	Tverrsnittsstudier	PRIMÆRSTUDIER/DATA: Nasjonale registre og statistikk, f.eks. www.ssb.no MEDLINE og/eller andre generelle ¹ databaser.
Hvordan kan man avgjøre om noen...?	Diagnose	Tverrsnittsstudier (med en gullstandard)	SYSTEMATISKE OVERSIKTER: Cochrane Library med CDSR, DARE og HTA supplert med søk i MEDLINE og/eller andre generelle databaser ² , PRIMÆRSTUDIER: MEDLINE og/eller andre generelle databaser
Hvordan vil det gå?	Prognose	Kohortstudier	SYSTEMATISKE OVERSIKTER OG PRIMÆRSTUDIER: MEDLINE og/eller andre generelle databaser
Hvorfor får noen...?	Årsak (etiologi)	Kohortstudier Kasus-kontrollstudier	SYSTEMATISKE OVERSIKTER: For spørsmål om bivirkninger ved en behandling: DARE og HTA (i Cochrane Library) supplert med søk i MEDLINE og/eller andre generelle databaser For vanlige årsaksspørsmål: MEDLINE og/eller andre generelle databaser PRIMÆRSTUDIER: MEDLINE og/eller andre generelle databaser
Hva kan man gjøre?	Effekt av tiltak	Randomiserte forsøk	SYSTEMATISKE OVERSIKTER: Cochrane Database of Systematic Reviews (i Cochrane Library) DARE og HTA (i Cochrane Library) supplert med søk i MEDLINE og/eller andre generelle databaser, Epistemonikos PRIMÆRSTUDIER: CENTRAL ³ (i Cochrane Library) MEDLINE og/eller andre generelle databaser
Hvordan oppleves det?	Erfaringer	Kvalitative studier	SYSTEMATISKE OVERSIKTER OG PRIMÆRSTUDIER: British Nursing Index ⁴ , CINAHL, MEDLINE, PsycINFO og/eller andre generelle databaser

¹ Med generelle databaser mener vi her databaser som ikke begrenser seg til spesielle studiedesign.

² DARE er en database i The Cochrane Library over kritisk vurderte oversikter som er laget utenfor Cochrane-samarbeidet. Den har etterslep, og derfor anbefales det å gjøre et tilleggssøk etter systematiske oversikter i MEDLINE og andre relevante databaser avgrenset til de siste årene, avhengig av hvor stort etterslepet er.

³ CENTRAL er en database som inneholder referanser til kontrollerte forsøk, i hovedsak randomiserte, og er den mest komplette oversikten over slike forsøk som finnes.

⁴ British Nursing Index er en database med et særlig fokus på sykepleiefaglig forskning, og den dekker kvalitativ forskning.

Bibliografiske databaser

En bibliografisk database registrerer opplysninger om tidsskriftartikler, rapporter, bøker, bokkapitler m.m. Slike opplysninger kalles referanser. Ingen database er komplett; de atskiller seg med hensyn til hvilke tidsskrifter som omfattes, typer av artikler som inkluderes, språk m.m. Eksempelvis er både MEDLINE og EMBASE internasjonale, men de har slagside mot tidsskrifter utgitt henholdsvis i USA og Europa. Det finnes også en rekke databaser for studier og oversikter publisert i ikke-vestlige land som ikke dekkes av MEDLINE og EMBASE, for eksempel LILACS (Sør-Amerika), African Trials Registry (Afrika) og IndMed (India). I tillegg til generelle medisinske og helsefaglige databaser, finnes det databaser som dekker et spesifikt fagområde. Et eksempel er PsycInfo, som dekker feltet psykisk helse.

Det finnes også databaser som inneholder spesielle studiedesign/metoder. Dette gjelder først og fremst kontrollerte studier og systematiske oversikter. Cochrane-samarbeidets register over kontrollerte studier (CENTRAL) er et eksempel.

Organisasjoners nettsider

Ikke alle publikasjoner indekseres i bibliografiske databaser, men de kan være publisert på nettsidene til organisasjonene som er ansvarlige for dem. Dette gjelder spesielt for sekundærlitteratur som retningslinjer, metodevarsler og HTA-rapporter og andre systematiske oversikter.

Grå litteratur og upubliserte studier

Resultater fra studier og oversikter presentert på konferanser og i rapporter utgitt av enkeltinstitusjoner publiseres ikke alltid i tidsskrift og blir derfor ofte ikke gjort tilgjengelig i databaser. Avhengig av tema kan det derfor være viktig å kontakte relevante organisasjoner. Det er ofte vanskelig og tidkrevende å finne slik litteratur, men det finnes databaser som registrerer "grått materiale", for eksempel OpenSIGLE: <http://opensigle.inist.fr/>. For norsk grå litteratur kan NORA (Norwegian Open Research Archives) være verdt å sjekke: <http://www.ub.uio.no/nora/search.html>. Det kan også ligge en del i OAIster, som samler litteratur fra mange åpne arkiv i verden: <http://www.oclc.org/oaister/>.

En del studier blir aldri publisert. De vanligste årsakene er at studien ikke viste tydelig effekt, at studien var liten eller at kommersielle interesser var involvert. Studier med positive resultater blir oftere publisert, og oftere på engelsk, enn mindre interessante funn. Å finne upubliserte studier er krevende. Vi forsøker å kontakte eksperter, industri og andre relevante aktører innenfor området, og leter på bestemte nettsteder. Det finnes også analytiske metoder som gjør det mulig å bedømme risiko for at det på et område forekommer selektiv publisering av studier og rapportering av «positive funn».

Pågående studier og oversikter

Den beste måten å imøtegå problemet med rapporteringskjevhet på, er gjennom en prospektiv registrering av alle pågående studier. Det finnes databaser for både studier og oversikter som er under arbeid. En oppdatert oversikt over slike kilder finnes i Vedlegg 1. Eksperter på feltet har ofte kunnskap om pågående studier. Det vil for de fleste forslag om oppsummeringer være nyttig å søke etter og liste opp pågående studier.

Gjennomgang av referanselister og siteringssøk

Selv systematiske søk i flere databaser vil ikke alltid identifisere all publisert og relevant litteratur. Det er ofte aktuelt å lete i referanselister i sentrale artikler. Man kan også søke via siterte arbeider (mulig via siteringsdatabasen Web of Science), og ta kontakt med forfatterne av publiserte studier og eksperter på feltet.

Håndsök

Håndsök, det vil si å lete manuelt gjennom årganger av spesielt relevante tidsskrifter, kan også være en aktuell framgangsmåte for å finne artikler som ikke finnes i databaser, eller som er dårlig indeksert. Dette er arbeidskrevende og gjøres i praksis sjelden i forbindelse med våre prosjekter.

LAGE SØKEALGORITME

Finne relevante søkeord

Det er viktig å finne aktuelle emneord, tekstord og fraser (sammenstillinger av ord) som fanger opp så mange som mulig av alle relevante referanser på et område (høy grad av sensitivitet/fullstendighet), og samtidig unngå ord som gir unødig mange irrelevante referanser (høy grad av spesifisitet), som vil føre til lav presisjon. Ideelt sett ønsker vi så høy fullstendighet, presisjon og spesifisitet som mulig. De søkene som utformes, er alltid et forsøk på å balansere disse faktorene.

De fleste helsefaglige databasene er engelskspråklige. En artikkel publisert i Tidsskrift for Den norske legeforening har for eksempel engelsk tittel og sammendrag (abstract) i MEDLINE selv om selve artikkelen er på norsk. I de fleste databasene må man derfor bruke engelske søkeord. Tabell 7 gir en oversikt over metoder for å finne relevante søkeord.

Tabell 7: Metoder for å finne relevante søkeord for et litteratursøk

Fagpersoner/eksperter	Bibliotekaren bør rådføre seg med fagpersoner med ekspertise på temaet for å finne relevante søkeord. I HTA-rapporter er det naturlig å konsultere utredningsgruppa hvis en slik gruppe er oppnevnt.
Databasenes emneordslister	Flere databaser har egne lister med emneord for å beskrive innholdet i artiklene (indeksere). MEDLINE/PubMed bruker Medical Subject Headings (MeSH), EMBASE bruker Emtree. Fordelen med faste emneord er synonymkontroll; i MEDLINE får artikler om brystkreft emneordet 'breast neoplasms' uansett om noen forfattere bruker 'breast cancer' eller andre synonymer i selve artikkelen.
Emneordlistenes definisjoner ('scope notes')	I systematiske søk holder det ikke bare å bruke emneord. Grunnen er at mange artikler ikke er tilstrekkelig indekserte. I mange tilfeller finnes heller ikke dekkende emneord, eller feltet er så nytt at det ennå ikke er etablert emneord. Det er derfor nødvendig å søke på tekstord i tillegg til emneord. Tekstord er ord i referansenes tittel eller sammendrag. Både selve emneordet og tilhørende synonymer søkes som tekstord. For å finne synonymer kan man gå til emneordenes definisjoner, i Ovid kalt 'scope notes'. Her finnes nyttig informasjon om emneordet ⁵ .
Ordbøker/oppslagsverk	Ordbøker/oppslagsverk er også nyttige kilder til søkeord. Eksempler er: Dorland's Medical Dictionary. Den gir gode definisjoner og forslag til synonymer og viser til relevante MeSH-termer. Ordboka finnes i nettsversjon, men er ikke gratis. MedlinePlus Medical Encyclopedia (http://www.nlm.nih.gov/medline_plus/encyclopedia.html) er et gratis oppslagsverk med definisjoner, synonymer og i noen tilfeller bilder og illustrasjoner.

Kombinere søkeord

Når vi har funnet alle tekstord og emneord i inkluderte databaser, må vi kombinere dem på riktig måte. Vi bryter spørsmålet (eller spørsmålene) ned i elementer knyttet til hvilke pasienter som er i fokus, hvilke intervensjoner og sammenlikninger som skal vurderes, hvilke utfall som er særlig viktige og hvilke design det er aktuelt å inkludere. Dette trinnet overlapper med det å etablere inklusjonskriterier.

De ulike kategoriene i spørsmålet bestemmer hvordan ordene kombineres. En generell regel er at vi kombinerer emneord og tekstord som beskriver hver kategori med OR, og til slutt alle kategorier med AND.

For flere databaser, deriblant MEDLINE og EMBASE, er det utarbeidet egne metodefiltre (søkefiltre) som fanger opp ulike typer studier, deriblant randomiserte studier og systematiske oversikter. Disse filtrene kan man legge til emnesøket for å fange opp relevante studier slik de er definert i inklusjonskriteriene. Se Vedlegg 1 for en oversikt over metodefilter for de ulike databasene.

Dette er et eksempel på oppsett av et søk:

1. Pasientord 1
2. Pasientord 2
3. 1 OR 2

⁵ I Ovids definisjoner finner man relevante synonymer under seksjonen 'used for'. Noen av ordene er ikke alltid synonymer. De kan også være en spesifisering av en mer generell term, der kun den generelle termen fungerer som emneord. Man bør gå gjennom definisjonene sammen med fagpersoner/eksperter.

4. Intervensjonsord 1
5. Intervensjonsord 2
6. 4 OR 5
7. 3 AND 6
8. Metodefilter
9. 7 AND 8

Her er et eksempel på utvikling av en søkealgoritme i MEDLINE⁶:

Vil antibiotikaprofylakse redusere sjansen for postoperative infeksjoner blant pasienter som hofteopereres?

Spørsmålet «brukket ned» i deler:

Populasjon	Pasienter som får operert inn nytt hoftede
Intervensjon	Forebyggende behandling med antibiotika (antimikrobiell profylakse)
Utfall	Postoperative infeksjoner
Studiedesign	Randomiserte kontrollerte forsøk

Identifisering av synonymer, stavevarianter og emneord til hver del:

	Emneord (MeSH)	Tekstord
Populasjon	Arthroplasty, Replacement, Hip	Hip replacement Arthroplasty (-ies)
Intervensjon	Antibiotic Prophylaxis	Antibiotic prophylaxis Antibiotics
Utfall	Surgical Wound Infection	Surgical wound infection(s) Postoperative infection(s)
Design	Randomized Controlled Trial	Randomized controlled trial Randomised controlled trial

Søkestrategi:

- | | |
|------------------------------------|----------------------------------|
| 1. Arthroplasty, Replacement, Hip/ | 9. surgical wound infection* |
| 2. hip replacement arthroplast* | 10. postoperative infection* |
| 3. total hip replacement | 11. 8 OR 9 OR 10 |
| 4. 1 OR 2 OR 3 | 12. 4 AND 7 AND 11 |
| 5. Antibiotic Prophylaxis/ | 13. Randomized Controlled Trial/ |
| 6. antibiotic* | 14. randomized controlled trial |
| 7. 5 OR 6 | 15. randomised controlled trial |
| 8. Surgical Wound Infection/ | 16. 13 OR 14 OR 15 |
| | 17. 12 AND 16 |

⁶ Dette er et forkortet, illustrerende eksempel. En komplett strategi ville inkludert flere synonymer for de ulike kategoriene, det er for eksempel langt flere måter å uttrykke postoperative komplikasjoner på enn de som er angitt her.

VURDERE RESULTATENE OG EVENTUELT REVIDERE SØKESTRATEGIENE

Å utvikle søkestrategier er en prosess med gjentakelser; den erfaring vi høster med innledende søk bruker vi til å forbedre strategiene. Hvis søket gir for mange irrelevante treff og/eller utelater relevante referanser, bør vi revidere søkestrategiene ved å legge til eller fjerne emneord og tekstord. Finner vi få referanser, bør vi vurdere å utvide søket til andre kilder. Noter om noen referanser er ekskludert på grunn av språk.

BEARBEIDE REFERANSENE VED HJELP AV REFERANSEHÅNDTERINGSVERKTØY

Søk etter studier for oversikter resulterer som oftest i store referansemengder. Den beste måten å håndtere store mengder referanser på, er å bruke et referansehåndteringsverktøy. Område for helsetjenester har tilgang til EndNote. Der kan vi:

- importere referanser fra de ulike databasene direkte inn i EndNote
- fjerne dubletter, dvs. referanser som finnes i mer enn én database, etter at alle referanser er importert
- holde orden på hvilke artikler som er i bestilling og hvilke som er mottatt
- legge inn egendefinerte felt der vi for eksempel angir hvilken database referansen ble funnet i

EndNote er integrert i MS Word, dvs. at vi kan legge inn referanser fra EndNote direkte i en tekst og automatisk generere en referanseliste i dokumentet.

LOGGFØRE OG RAPPORTERE SØK

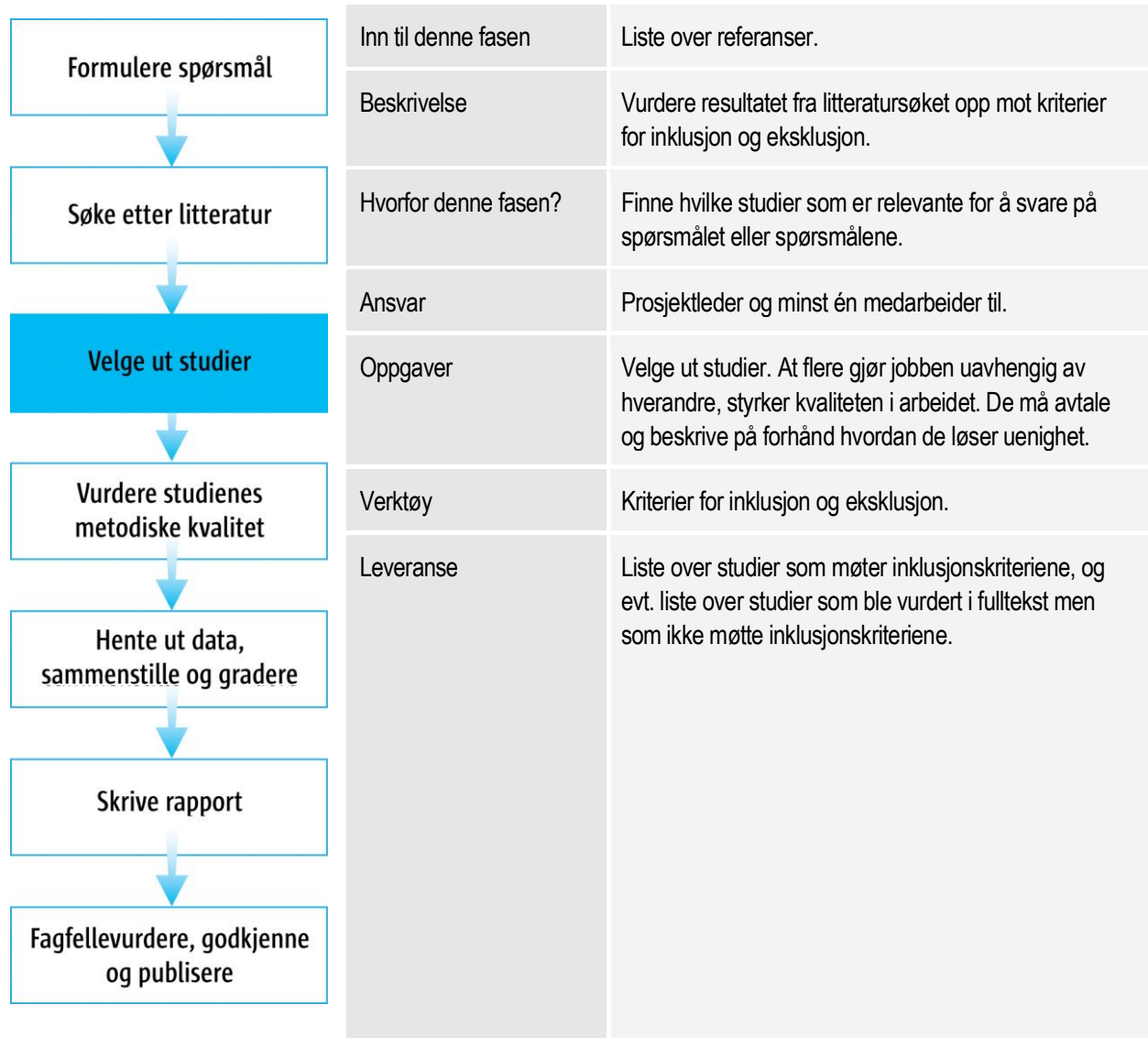
Det er viktig å loggføre valgene vi gjør underveis i søkeprosessen for å dokumentere valg av kilder, søkeord og kombinasjonen av søkeord. Dette er ikke minst viktig hvis søket skal oppdateres på et senere tidspunkt. Alle søkene skal alltid vises i vedlegg til oversiktene våre.

OPPDATERE SØK

I mange databaser er det mulig å lagre søkestrategien og gjenta søket på et senere tidspunkt. Blant annet har databasene som er tilgjengelige via Ovid og Cochrane Library, denne muligheten. Ovid tillater også å avgrense et oppdatert søk fra eksakt dato/uke for forrige søk. Framgangsmåten varierer fra database til database.

Tid mellom dato for søk og dato for publisering av oversikten bør helst ikke overskride 8-12 måneder. Dette er for å redusere sjansen for at nye, viktige studier som kan ha betydning for resultat i oversikten er publisert etter søket (og dermed ikke tatt med i oversikten). Hvis det finnes tungtveiende grunner for at tidsrammen ikke opprettholdes, må det redegjøres for dette i oversikten.

5 Velge ut studier



ER ARTIKKELSAMMENDRAGET RELEVANT?

For å finne relevante studier skal minst to personer uavhengig av hverandre vurdere referansene fra litteratursøket opp mot kriterier for inklusjon og eksklusjon. De bør helst gjøre dette først ut fra sammendrag (ikke bare tittel).

Referanser begge personene mener er irrelevante, tas ikke med. De som vurderer skal hente inn fulltekstversjon av referanser når de er:

- enige om at referansen kan være relevant
- uenige eller usikre om hvorvidt referansen er relevant

Reproduserbarheten av denne prosessen bør sjekkes i starten av utvelgelsen; det kan være nødvendig å lage mer eksplisitte kriterier. Uenighet kan skyldes at man overser informasjon eller har ulik tolkning av informasjonen. Uenighet bør diskuteres og helst løses i fellesskap på grunnlag av det som står i prosjektplanen.

ER SELVE PUBLIKASJONEN RELEVANT?

To personer (et lesepar) leser publikasjonene (som oftest dreier det seg om artikler) i fulltekst uavhengig av hverandre og vurdere fulltekstene opp mot kriterier for inklusjon og eksklusjon.

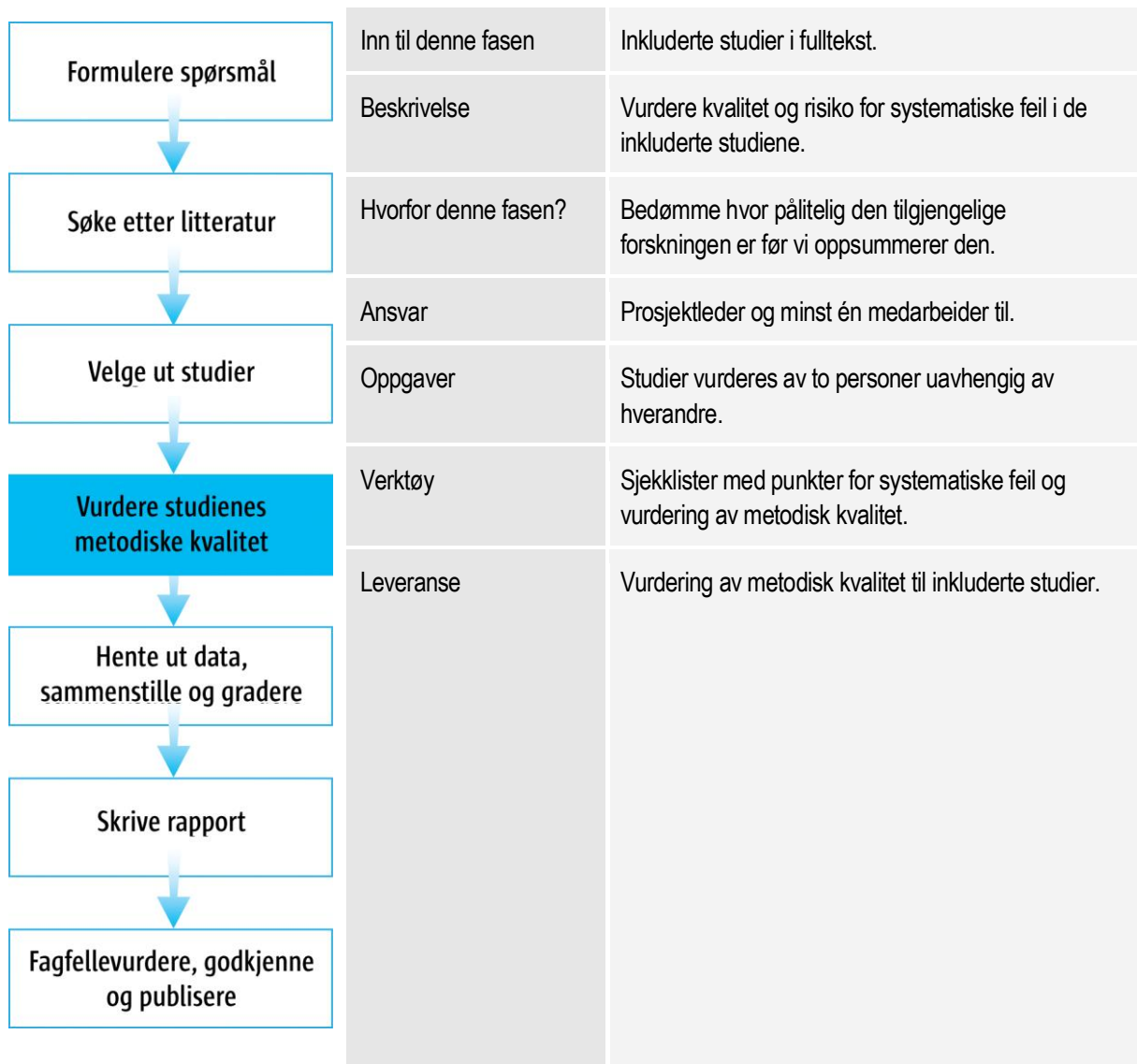
Det er opplagt et innslag av subjektivitet i utvelgelsen av studier, selv om inklusjonskriteriene er laget så eksplisitte som mulig. Dersom det er tid og anledning, er det en fordel at prosessen gjennomføres av et panel av spesielt metodekyndige personer og av eksperter på feltet. Når vi arbeider sammen med utredningsgrupper, kan de derfor trekkes med i utvelgelsen.

På dette tidspunktet lages det som regel en tabell over ekskluderte studier der vi beskriver årsaken til at studien ikke er tatt med. Den endelige rapporten skal inneholde et flytdiagram som dokumenterer utvelgelsen; hvor mange referanser som ble identifisert ved søket, og hvor mange publikasjoner som ble inkludert og ekskludert ved vurdering av sammendrag og gjennomgang av publikasjoner i fulltekst.

Viktige trinn i seleksjonen av studier

- I utvelgelsen av studier er det viktig å bestrebe seg på å unngå seleksjonsskjevheter. Kriteriene skal logisk følge av spørsmålet oversikten skal besvare, og de bør prøves ut for å sjekke om studier identifiseres på en pålitelig måte.
- Utvelgelsen foregår i stadier der man først vurderer titler og sammendrag, så innhenter mulig relevante publikasjoner i fulltekst og etter gjennomlesing tar med studier som oppfyller kriteriene for inklusjon.
- Vi gjennomfører parallelle vurderinger som skjer uavhengig av hverandre. På forhånd blir vi enige om hvordan vi løser uenighet.
- Vi skal dokumentere utvelgelsen og begrunne inklusjon og eksklusjon.

6 Vurdering av inkluderte studier



VURDERING AV INKLUDERTE STUDIER

Etter at alle identifiserte publikasjoner er gjennomgått i forhold til inklusjonskriteriene og vurdert for relevans, vurderer vi hvordan de inkluderte studiene eller systematiske oversiktene er utført. Dette hjelper oss til å finne ut i hvilken grad vi kan stole på resultatene.

Tolkningen av en studie avhenger både av intern og ekstern validitet (gyldighet). Måten en undersøkelse om effekt er gjennomført på, påvirker vår tillit til at resultatet er kommet fram gjennom en god tilnærming til den sanne, underliggende effekten (intern validitet). Ekstern validitet er avhengig av om de personene, intervusjonene, sammenlikningene og utfallsmålene som er inkludert i studien, er tilstrekkelig like det som er aktuelt i ordinær virksomhet i helsetjenesten til at vi har tillit til at resultatene er overførbare til vår virkelighet.

Når vi vurderer den interne validiteten, ser vi på om forskningsspørsmålet er besvart på en måte som gjør at resultatene er mest mulig fri for systematiske feilkilder (bias). Systematiske feil er metodefeil som gir resultater som (systematisk) avviker fra «sannheten», f.eks den sanne underliggende effekten av et tiltak. Systematiske feil kan ved spørsmål om effekt føre til overestimering eller underestimering av den sanne effekten av et tiltak. Feilene kan være trivielle, eller de kan være så store at et tilsynelatende positivt funn helt og holdent skyldes en systematisk feilkilde. Hver potensielle kilde for systematiske feil kan ha forskjellige virkninger i forskjellige studier. For eksempel er skjult fordeling til tiltaks- og kontrollgruppe og blinding av stor viktighet for subjektive utfallsmål, men mindre viktig for objektive utfallsmål (10).

Alle som driver forskning bør bruke metoder som best mulig beskytter mot ulike typer av systematiske feilkilder. Fordi det ikke alltid er tilfelle, forsøker vi å gjøre en vurdering av hvilke metodiske mangler/ mulige feilkilder som er i studien eller den systematiske oversikten, og i hvilken grad manglene får innvirkning på vår tillit til resultatene. Et visst innslag av tilfeldige feil kan vi aldri beskytte oss mot, og forskere tar høyde for dette ved sammenstilling og vurdering av tillitt til resultatene.

I område for helsetjenesten i Folkehelseinstituttet bruker vi to ulike tilnæringsmåter når vi vurderer de studiene eller systematiske oversiktene vi inkluderer. For systematiske oversikter vurderer vi den metodiske kvaliteten av oversikten. For primærstudier vurderer vi risiko for systematiske feil i resultatene. Begrepet risiko for systematiske feil er forskjellig fra begrepet kvalitet. Kvaliteten av en studie bestemmes av i hvilken grad forskerne utførte forskningsprosjektet i samsvar med en akseptert standard for hvordan studier bør utføres. En studie kan imidlertid være utført i samsvar med alle standarder, men likevel kan det være betydelig risiko for systematiske feil i resultatet. For

eksempel er det i noen situasjoner umulig å gjennomføre blinding av helsepersonell og deltakere. Da blir det urimelig å si at det er en dårlig studie, men manglende blinding kan uansett utgjøre en høy risiko for systematiske feil i resultatet som vi da trekker for i vurderingen av systematiske feil. Omvendt kan det også være kvalitetsmessige svakheter i studien som ikke nødvendigvis fører til risiko for systematiske feil i resultatet.

Å VURDERE OVERSIKTERS METODISKE KVALITET

Når vi vurderer den metodiske kvaliteten til systematiske oversikter benytter vi en sjekkliste som belyser ulike sider av hvordan den systematiske oversikten er utført. Vi ønsker å forsikre oss om at alle relevante studier er identifisert og inkludert, og at forfatterne har gjort en vurdering av kvalitet eller eventuelle systematiske feil i de studiene de har inkludert. Sjekkliste for vurdering av systematiske oversikters metodiske kvalitet finnes på FHI's nettsider.

Uavhengig av hverandre vurderer to personer oversiktens metodiske kvalitet ved hjelp av sjekklisten. Oversiktene klassifiseres som høy, middels eller mangelfull mht metodisk kvalitet avhengig av hvor mange og hvilke av kriteriene i sjekklisten som er oppfylt.

I oppsummering av systematiske oversikter vurderer vi ikke selv kvaliteten eller risiko for systematiske feil i primærstudiene som er inkludert i oversikten. Her bruker vi oversiktforfatterens vurderinger.

Å VURDERE SYSTEMATISKE SKJEVHETER I PRIMÆRSTUDIER

For vurdering av systematiske skjevheter i primærstudier fins det flere ulike typer sjekklister for ulike typer studiedesign. For vurdering av metodiske skjevheter i randomiserte kontrollerte studier benytter vi et verktøy som er beskrevet i The Cochrane Handbook, kalt 'Risk of bias' (2). Dette er et sett av kriterier som brukes for å gjøre en kritisk vurdering av kjente kilder for systematiske feil. Disse kriteriene er bl.a. generering av randomiseringssekvens, skjult fordeling til grupper, blinding og vurdering av ufullstendige data eller manglende rapportering om utfallsmålene. Mangelfull blinding og frafall ved oppfølging kan gi forskjellig risiko for systematisk skjevheter for ulike typer utfall. Risikoen for systematiske feil bør derfor vurderes for hvert enkelt utfall eller grupper av utfall. Dette forutsetter en betydelig grad av innsikt i forskningsmetode, i tillegg til bruk av skjønn. To personer gjør denne vurderingen uavhengig av hverandre.

Slik kan vi vurdere informasjon om risiko for systematiske skjevheter for det enkelte utfallsmålet som forklaring på at forskjellige studier eventuelt har ulike resultater. Vi bruker også denne informasjonen for å kunne legge større vekt på studier med lavere risiko for systematiske feil, for eksempel ved å utføre sensitivitetsanalyser kun for studier med lav risiko for systematiske skjevheter.

I hvert prosjekt er det nødvendig å vurdere om det er andre forhold enn de som er nevnt i sjekklistene, som vi bør vurdere i tillegg. Sjekkliste vi bruker for å vurdere risiko for systematiske skjevheter i primærstudier, finnes på FHIs nettsider.

Cochranes klassifikasjon av systematiske skjevheter (bias) (2) i randomiserte kontrollerte studier er:

Table 8.5.a: The Cochrane Collaboration's tool for assessing risk of bias

Domain	Support for judgement	Review authors' judgement
<i>Selection bias.</i>		
Random sequence generation.	Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.	Selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence.
Allocation concealment.	Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.	Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment.
<i>Performance bias.</i>		
Blinding of participants and personnel <i>Assessments should be made for each main outcome (or class of outcomes).</i>	Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Performance bias due to knowledge of the allocated interventions by participants and personnel during the study.
<i>Detection bias.</i>		
Blinding of outcome assessment <i>Assessments should be made for each main outcome (or class of outcomes).</i>	Describe all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Detection bias due to knowledge of the allocated interventions by outcome assessors.
<i>Attrition bias.</i>		
Incomplete outcome data <i>Assessments should be made for each main outcome (or class of outcomes).</i>	Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.	Attrition bias due to amount, nature or handling of incomplete outcome data.
<i>Reporting bias.</i>		
Selective reporting.	State how the possibility of selective outcome reporting was examined by the review authors, and what was found.	Reporting bias due to selective outcome reporting.
<i>Other bias.</i>		
Other sources of bias.	State any important concerns about bias not addressed in the other domains in the tool. If particular questions/entries were pre-specified in the review's protocol, responses should be provided for each question/entry.	Bias due to problems not covered elsewhere in the table.

SAMLET VURDERING AV RISIKO FOR SYSTEMATISKE SKJEVHETER

Hensikten med å vurdere risiko for systematiske skjevheter for hvert utfallsmål eller gruppe av utfallsmål, er å angi i hvilken grad vi har tillit til resultatene. Etter at vi ved hjelp av sjekklisten har undersøkt i hvilken grad kriteriene er oppfylt, bør vi gjøre en samlet vurdering av risiko for systematisk skjevheter i utfallene på tvers av studiene. Vi bruker kategoriene liten, uklar eller høy risiko

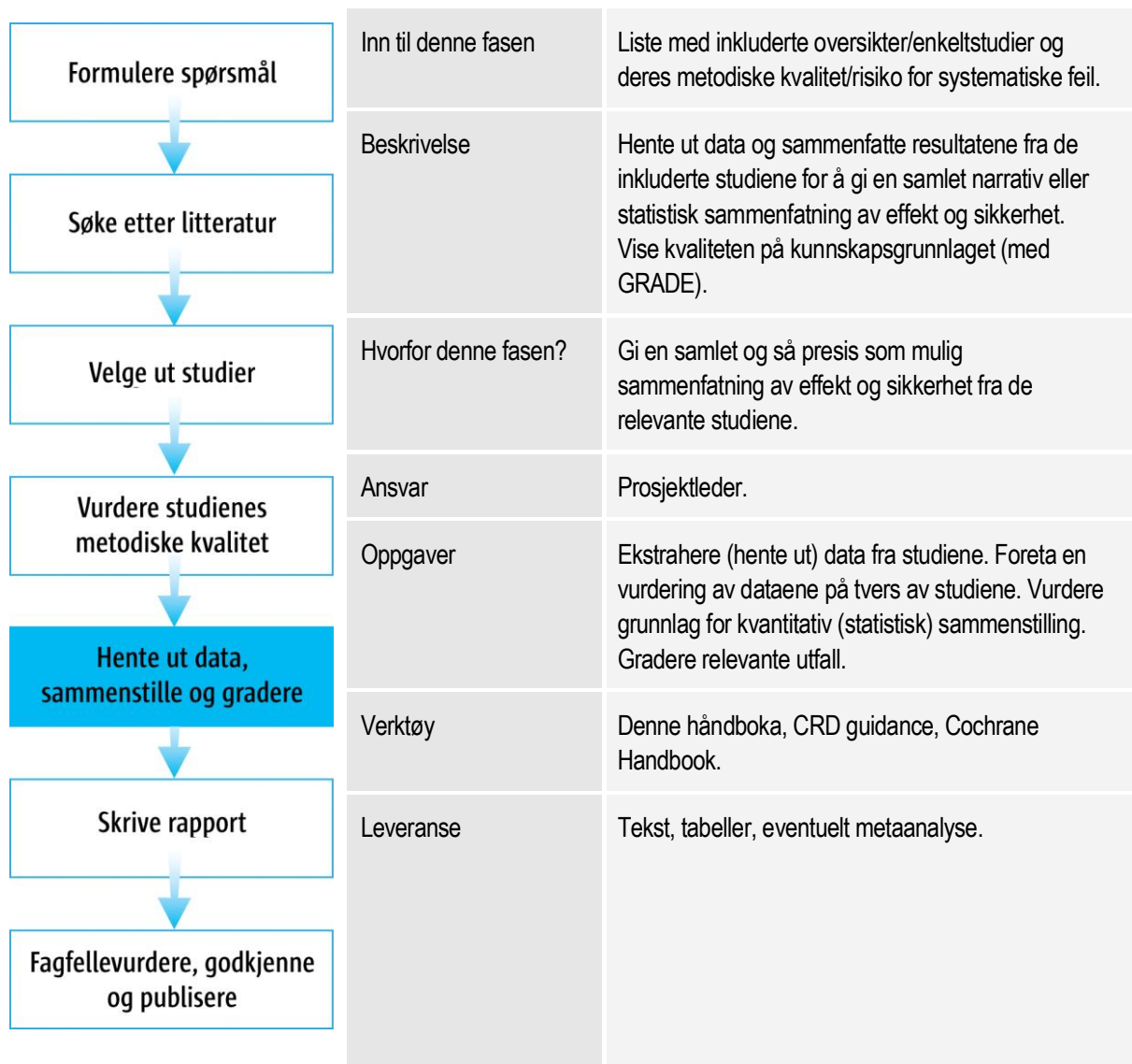
for systematiske skjevheter. I Cochrane Handbook (2) eksemplifiseres dette som beskrevet nedenfor.

Mulig tilnærming for å oppsummere vurderinger av risiko for systematiske skjevheter (bias) for hvert utfall på tvers av studier (2):

Table 8.7.a: Possible approach for *summary assessments* of the risk of bias for each important outcome (across domains) within and across studies

Risk of bias	Interpretation	Within a study	Across studies
Low risk of bias.	Plausible bias unlikely to seriously alter the results.	Low risk of bias for all key domains.	Most information is from studies at low risk of bias.
Unclear risk of bias.	Plausible bias that raises some doubt about the results.	Unclear risk of bias for one or more key domains.	Most information is from studies at low or unclear risk of bias.
High risk of bias.	Plausible bias that seriously weakens confidence in the results.	High risk of bias for one or more key domains.	The proportion of information from studies at high risk of bias is sufficient to affect the interpretation of results.

7 Hente ut data, sammenstille og gradere



HENTE UT DATA

Dataekstraksjon vil si å hente ut relevant informasjon fra de inkluderte studiene eller oversiktene. Prosjektplanen spesifiserer hvilke data som er relevante, og som ekstraheres fra den enkelte studien eller oversikten. Det kan være nyttig å

bruke et dataekstraksjonsskjema. Dataekstraksjon gjennomføres av to personer uavhengig av hverandre, eller av én person og sjekkes så av en annen. Uenighet løses ved diskusjon eller ved å konsultere en tredjeperson.

Viktige elementer i et skjema for dataekstraksjon fra effektstudier

- Tittel, forfattere, detaljer om publikasjonen
- Plass til å notere spørsmål og annet det er viktig å huske på
- Endelig verifisering av at publikasjonen oppfyller inklusjonskriteriene og bør inkluderes

Kjennetegn ved primærstudier

- Populasjonen (antall inkludert i hver gruppe, alder, kjønn, diagnose, alvorlighetsgrad, komorbiditeter)
- Hvor og når studien ble utført (årstall, land, sykehus, utenfor sykehus osv.)
- Intervensjonen(e)
- Sammenlikning(er)

Kjennetegn ved systematiske oversikter

- Kvaliteten på den systematiske oversikten
- Dato for litteratursøket
- Inkluderte sammenlikninger

Utfallsmål og resultater

- Utfallsmål (som spesifisert i prosjektplanen)
 - Diskrete data: Antall hendelser og antall personer i gruppene (risk ratio (RR), odds ratio (OR)), p-verdier og konfidensintervall (KI)
 - Kontinuerlige data: Gjennomsnitt, standardavvik, andre statistiske mål
 - Data som er målt over tid («overlevelsesdata»)
- Oppfølgingstid og målemetode
- Justerte data der de er tilgjengelige (og informasjon om hvilke faktorer det er justert for)
- Frafall
- Hvordan manglende data er håndtert

BRUK AV DATA FRA SYSTEMATISKE OVERSIKTER

I en systematisk oversikt er resultatene fra studiene allerede sammenstilt. Oppgaven til forskerne er da å vurdere og sammenfatte relevante data fra én eller flere systematiske oversikter. I dette er det flere utfordringer.

Ved inklusjon og gjennomgang av systematiske oversikter er det fire forhold som er avgjørende:

1. Er den systematiske oversikten relevant for problemstillingen (vår PICO)?
 - a. PICO er fullstendig overlappende
 - b. PICO er delvis overlappende
2. Er dette en systematisk oversikt av tilstrekkelig kvalitet (primært høy, ev. middels) til at vi kan stole på resultatene (jfr. sjekkliste for systematiske oversikter)
3. Vil resultatene være overførbare til vår setting (norsk helsetjeneste)?
4. Er søket nytt nok til å vurderes som oppdatert? Eller trenger denne oversikten oppdatering for å være nyttig?

Ved gjennomgang av resultatene er det vår PICO som er førende for hvilke data som skal formidles fra den systematiske oversikten. Dersom det mangler data for enkelte endepunkter (for eksempel om sikkerhet), bør det spesifiseres.

Informasjon om de systematiske oversiktene presenteres i en tabell som viser karakteristika for inkluderte systematiske oversikter og kvalitetsvurdering. Resultatene presenteres i tekst og ved GRADE-tabeller (se senere i kapittelet).

Mer om sammenstilling av data fra systematiske oversikter fins i Cochrane Handbook, kapittel 22 (2).

BRUK AV DATA FRA PRIMÆRSTUDIER

Når vi bruker data fra primærstudier, må vi vurdere og sammenstille resultatene for hver sammenlikning. Dersom den systematiske oversikten skal inkludere flere studiedesign, må disse sorteres hver for seg. Deretter sammenfattes resultatene for hvert av de relevante utfallsmålene. Dette kan gjøres ved en deskriptiv, ikke-kvantitativ syntese, eller når det ligger til rette for det, en kvantitativ syntese ved bruk av statistiske teknikker (metaanalyse). I begge tilfeller er det viktig å vurdere hvorvidt effektene er konsistente på tvers av alle de inkluderte studiene, både de som kan inngå i en metaanalyse og de som eventuelt ikke har data som kan benyttes i metaanalysen.

Dersom det påvises inkonsistens, må studiene vurderes for mulige årsaker til dette. Resultatene presenteres i figurer (forest plot) og GRADE-tabeller (se senere i kapittelet).

BESKRIVELSE AV STUDIENE OG RESULTATENE

I de fleste typer oversikter er det vanlig at informasjon om de inkluderte studiene presenteres i en tabell som viser studiekarakteristika og en tabell som viser risiko for bias. I tillegg oppsummeres og beskrives resultatene fra studiene.

Narrativ syntese / beskrivende syntese

Dersom studiene er for ulike (klinisk eller metodisk heterogenitet) til at resultatene kan sammenfattes i en metaanalyse, bør de beskrives hver for seg samt sammenstilles ved en «beskrivende» syntese. Vi bør unngå å oppsummere studiene ved å telle antall studier som viser en effekt og antall som ikke gjør det for å gi en samlet konklusjon, da dette ofte blir misvisende fordi studiene kan ha ulik kvalitet og størrelse.

Målet med en beskrivende syntese er å gi en beskrivelse av de samlede resultatene, *retningen* og *størrelsen* på effektestimaterne. Dette gjøres ved å vurdere sammenheng mellom resultatene innen og mellom studiene, og gi et inntrykk av hvor konsistente og robuste resultatene er. Prosessen forutsetter subjektive vurderinger, og det er derfor viktig å ha klare kriterier for hvilke elementer som vektlegges i vurderingene.

Resultatene presenteres for de effektmålene som er beskrevet i prosjektplanen. Beskrivelsen bør gjøre rede for hvor mange studier av de totalt inkluderte for en sammenlikning som har data på effekt og sikkerhet.

Effekttestimat kan måles og presenteres på mange ulike måter og med en rekke metoder. Det er ønskelig at de metodene som benyttes er validerte, også for den populasjonen de anvendes på.

Hvilke effektestimater som vektlegges, skal være beskrevet i prosjektplanen.

Dikotome utfall

Dikotome utfall brukes når man klart kan si at en hendelse har skjedd eller ikke (for eksempel død/ikke død). Effektestimater presenteres som absolutte verdier (antall hendelser, absolutt risikoforskjell eller number-needed-to-treat (NNT)), og som relative verdier (risk ratio, odds ratio, relativ risikoreduksjon, prosent).

Kontinuerlige data

Kontinuerlige data brukes når effektestimater måles på en skala. Effektestimater kan presenteres som absolutte eller relative gjennomsnittlige forskjeller mellom gruppene. Ofte vil en erfare at studier har rapportert samme utfall, men benytter

ulike skalaer for å måle det. Da er en mye brukt teknikk å beregne standardisert gjennomsnittlig forskjell (standardized mean difference). Det kan være nyttig å beregne det samlede effektestimater tilbake til en av skalaene for å vise hva resultatene kan bety i praksis.

Tidskurver (overlevelseskurver)

Tidskurver brukes for å analysere tid til en hendelse, men der ikke alle de inkluderte nødvendigvis opplever hendelsen. Et eksempel er overlevelseskurver. Vi bør tilstrebe å rapportere konfidensintervall da det er begrensninger ved kun å rapportere p-verdier.

Det står mer om ulike effektestimater i Cochrane Handbook, kapittel 9.2 (2).

Metaanalyse

Ofte vil ulike studier presentere data på forskjellige måter; det kan være ulike skalaer, ulike tidspunkt for måling eller ulike terskler med mer. Det kan også være en rekke andre forskjeller mellom studiene (populasjon, type intervensjon eller sammenlikning). Derfor er det nødvendig å vurdere hvilke data som er like nok til å kategoriseres og beskrives sammen, og hvilke data som kan inngå i en metaanalyse. Det er som regel hensiktsmessig å konsultere med en statistiker ved utførelse av metaanalyser.

Metaanalyse uatarbeides dersom det er beskrevet i prosjektplanen og det er et tilstrekkelig homogent datagrunnlag. En slik syntese forutsetter ikke bare at man i primærstudiene har undersøkt samme problemstilling på sammenliknbare utvalg av deltakere, men også at man har målt det samme med sammenliknbar metode. Når vi gjør metaanalyser, følger vi hovedsakelig metodene som er beskrevet i Cochrane Handbook, kapittel 9.4 (2).

Det er tre spørsmål som må avklares:

- Hvilke sammenlikninger skal gjøres?
- Hvilke utfallsmål skal brukes?
- Hvilke effektmål skal brukes?

Disse forholdene bør man ha tenkt over og beskrevet i prosjektplanen. Det kan imidlertid være nødvendig å modifisere planen når det blir klart hva slags data de enkelte studiene rommer. Eventuelle modifiseringer av planen må dokumenteres.

I metaanalysen tar vi effektestimater fra hver enkelt studie. Dette vektet, som regel basert på hvor mange hendelser og deltakere hver studie hadde, og et samlet effektestimater beregnes for hvert endepunkt. Det gjøres med dataprogrammer, som også beregner hvor presist estimatet er (dvs. etablerer et

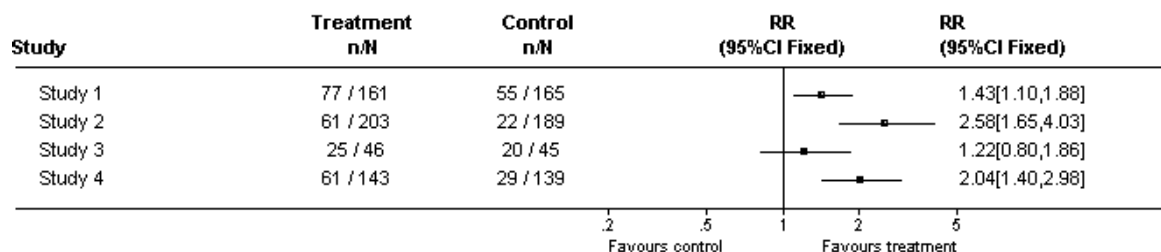
konfidensintervall). Resultatene presenteres i et såkalt «forest plot» som vist nedenfor. Firkantene viser effektestimater fra de enkelte studiene, og de horisontale linjene viser konfidensintervallene. Diamanten nederst viser det samlede effektestimateret fra metaanalysen, og bredden av den viser konfidensintervallet. Den vertikale linjen er «the line of no difference». Ligger effektestimateret på den, kan vi ikke se bort fra at det kan gå likedan med personene som fikk et bestemt tiltak som med personene i kontrollgruppa.

Fordi studiene som kombineres er forskjellige, vil det alltid være noe variasjon eller heterogenitet mellom dem. Heterogenitet kan være *klinisk* (forskjeller mellom deltakerne, tiltakene eller utfallsmålene), *metodisk* (forskjell i studiedesign eller risiko for skjevheter (bias) mellom studiene eller *tilfeldig*. Effektestimater fra studiene kan kombineres med en "fixed effect"-modell der det antas at de enkelte studiene kan sees på som deler av én stor studie, og en "random effect"-modell der man åpner for at det kan være systematiske forskjeller mellom enkeltstudiene. De fleste av våre sammenstillinger egner seg for "random effect".

Nedenfor gir vi eksempler på metaanalyser.

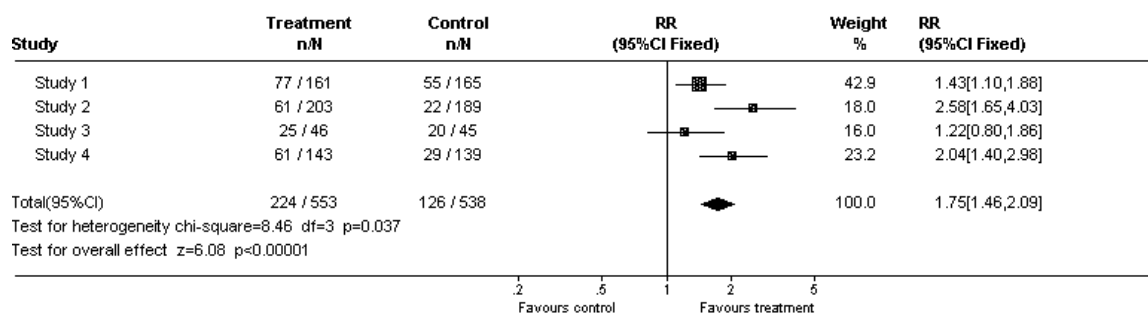
Effects of four trials included in a systematic review⁷

a) Presented without metaanalysis

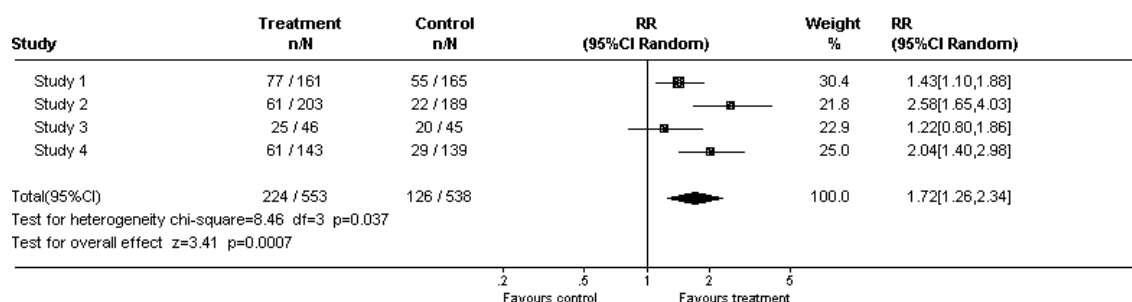


⁷ Trykket med tillatelse. Hentet fra Undertaking Systematic Reviews of Research on Effectiveness. CRD's Guidance for those Carrying Out or Commissioning Reviews. CRD Report 4, Centre for Reviews & Dissemination, University of York, 2001.

b) Presented with metaanalysis (fixed effect model)



c) Presented with metaanalysis (random effect model)



Figur «Forest plot» av tre ulike måter å sammenstille data på.

GRADERING

Når vi sitter med en samling av relevante studier som har informasjon om en bestemt sammenlikning og et utfall (f.eks. om effekten av et bestemt tiltak mot smerte sammenliknet med et annet tiltak eller ikke noe tiltak), skal vi vurdere hvilken tillit vi har til resultatene. Er det grunn til å tro at dette samlet sett er et robust effekttestimat, som neppe vil bli påvirket selv om det kommer nye studier? Spriker resultatene i en grad som gjør oss usikre på hva den sanne, underliggende effekten egentlig er? Og sett i forhold til vårt opprinnelige spørsmål: Treffer studiene helt med hensyn på type tiltak, personer det er prøvd ut på og utfall som er målt?

I område for helsetjenesten i Folkehelseinstituttet bruker vi GRADE (Grading of Recommendations Assessment, Development and Evaluation), som definerer tilliten til den samlede dokumentasjonen slik:

GRADE	Definisjon	Symbol

Høy tillit	Vi har stor tillit til at effektestimaten ligger nær den sanne effekten	⊕⊕⊕⊕
Middels tillit	Vi har middels tillit til effektestimaten: Det ligger sannsynligvis nær den sanne effekten, men det er også en mulighet for at den kan være forskjellig	⊕⊕⊕□
Lav tillit	Vi har begrenset tillit til effektestimaten: Den sanne effekten kan være vesentlig ulik effektestimaten	⊕⊕□ □
Svært lav tillit	Vi har svært liten tillit til at effektestimaten ligger nær den sanne effekten	⊕□□ □

GRADE er en metode for å vurdere tilliten til dokumentasjonen og styrken på anbefalingene. Kunnskapsoppsummeringene våre gir ikke anbefalinger og bruker derfor bare GRADE til å vurdere tilliten til dokumentasjonen for hvert utfall. GRADE kjennetegnes ved transparens. Vi viser fram vurderingene som vi har gjort i fotnoter til GRADE-tabellene (11).

Utgangspunktet for bruk av GRADE er en systematisk oversikt av høy metodisk kvalitet over spørsmålet vi søker å besvare. Dette vil enten være oversikten vi er i ferd med å lage eller en oversikt laget av andre. Spørsmålet definerer hvilke virkemidler eller intervensjoner som sammenliknes (f.eks. massasje versus varmebehandling), hvilke mennesker intervensjonen skal testes på (f.eks. voksne med akutte korsryggsmerter) og hvilke utfall som er av interesse (f.eks. positive utfall som mindre smerte, og negative utfall som bivirkninger).

I GRADE-systemet utvikles det en evidensprofil for det spesifikke spørsmålet som skal belyses. En GRADE-profil viser tydelig hvilke vurderinger som er utført under graderingen av kvaliteten på dokumentasjonen, for hvert utfall som er definert. I tillegg presenteres en oppsummeringstabell. Metoden innebærer blant annet at en studie som i utgangspunktet starter på lav kvalitet, kan oppgraderes, og at en studie som starter på høy kvalitet, kan nedgraderes.

Kriterier for vurdering av tillit med GRADE

Studiedesign er utgangspunkt for vurderingen av utfallene. Det deles inn i to hovedkategorier: randomiserte kontrollerte studier og observasjonsstudier. Randomiserte kontrollerte studier starter på høy, og observasjonsstudier starter på lav. Vi vurderer åtte kriterier: fem som kan resultere i nedgradering, og tre som kan føre til oppgradering.

Studiekvalitet handler om planlegging og utføring av studien. For hvert av utfallene må vi vurdere risiko for systematiske skjevheter. Dette er allerede utført som del av den systematiske oversikten. Vi noterer hvorfor tilliten eventuelt blir nedgradert.

Konsistens eller samsvar mellom studiene handler om hvor like effektestimater for utfallet er mellom de forskjellige studiene. Heterogenitet (manglende samsvar) kan vise seg ved at resultatene fra forskjellige studier peker i forskjellige retninger, eller at det er stor forskjell i effektstørrelsen mellom ulike studier. Er det uforklarlig heterogenitet i resultatene, nedgraderes vår tillit til dette utfallet. Hvis heterogenitet kan forklares med forskjeller i pasientkarakteristika, diagnose eller risikogruppe, bør vi vurdere om vi bør lage en egen evidensprofil for hver av disse gruppene.

Direkthet handler om hvor like studiedeltakere, intervensjon, sammenlikning og utfallsmål i de inkluderte studiene er i forhold til spørsmålet som stilles:

- Har vi grunn til å tro at studiedeltakerne ikke er representative for pasientene dokumentasjonen skal brukes på, for eksempel hvis de er sykere eller eldre, nedgraderes vår tillit for dette utfallet.
- Hvis intervensjonen er vesentlig forskjellig fra den som spørsmålet egentlig beskriver, for eksempel med hensyn til type legemiddel, dosering eller behandlingsintensitet, nedgraderes vi for dette utfallet.
- Er utfallet viktig, riktig målt og direkte relevant til spørsmålet? Bruk av surrogatutfall kan gi misvisende resultater.
- En annen form for mangel på direkthet er når det mangler direkte sammenlikning mellom de to intervensjonene som vurderes: for eksempel om man vil sammenlikne effekten av A og B, men bare finner A sammenliknet med placebo og B sammenliknet med placebo.

Presisjon handler om hvor presise resultatene er, hvor mye data vi har og hvor stor usikkerheten i resultatene er. Hvis det kun er én studie eller studien(e) kun inkluderer noen få hendelser, og estimatet dermed har brede konfidensintervaller, nedgraderes vår tillit til dette utfallet.

Rapporteringsskjevheter handler om både publiseringsskjevheter og skjevheter i hvilke utfall som rapporteres. Dersom vi har grunn til å tro at det foreligger rapporteringsskjevhet, nedgraderer vi for det aktuelle utfallet. Dersom vi har grunn til å tro at det foreligger publiseringsskjevhet, nedgraderer vi.

De tre kriteriene i GRADE-systemet for å oppgradere tilliten til dokumentasjonen gjelder kun der det er flere samsvarende observasjonsstudier som ikke er nedgradert:

- **Sterke eller veldig sterke assosiasjoner/sammenhenger** mellom intervensjon og utfall.
- **Store eller veldig store dose–responseffekter** (jo sterkere «dose» av intervensjonen, jo sterkere respons).
- Der alle **forvekslingsfaktorer** ville ha redusert effekten.

Av tabellen nedenfor framgår det hvordan vi bruker de aktuelle kriteriene til å vurdere kvaliteten på dokumentasjonen for viktige utfall i en systematisk oversikt.

Tabell 8: Kriterier for vurdering av kvaliteten på dokumentasjonen for hvert enkelt utfall ved GRADE

Vår tillit til dokumentasjonen	Studiedesign	Nedgrader ved	Oppgrader ved*
Høy	Randomisert kontrollert forsøk	Begrensninger ved studiekvaliteten (risiko for systematiske feil)	Sterk sammenheng +1 Sterk +2 Veldig sterk
Middels		-1 Alvorlig -2 Veldig alvorlig	Dose-responseeffekt +1 holdepunkter for dose-responseeffekt
Lav	Observasjonsstudier	Mangel på samsvar -1 Alvorlig -2 Veldig alvorlig	Forvekslingsfaktorer +1 Alle kjente forvekslingsfaktorer ville ha redusert effekten
Svært lav		Mangel på direktehet -1 Alvorlig -2 Veldig alvorlig eller flere faktorer Mangel på presisjon -1 Alvorlig -2 Veldig alvorlig Rapporteringsskjevheter -1 Sannsynlig -2 Veldig sannsynlig	

*Forutsetter at det er flere samsvarende observasjonsstudier som ikke er nedgradert.

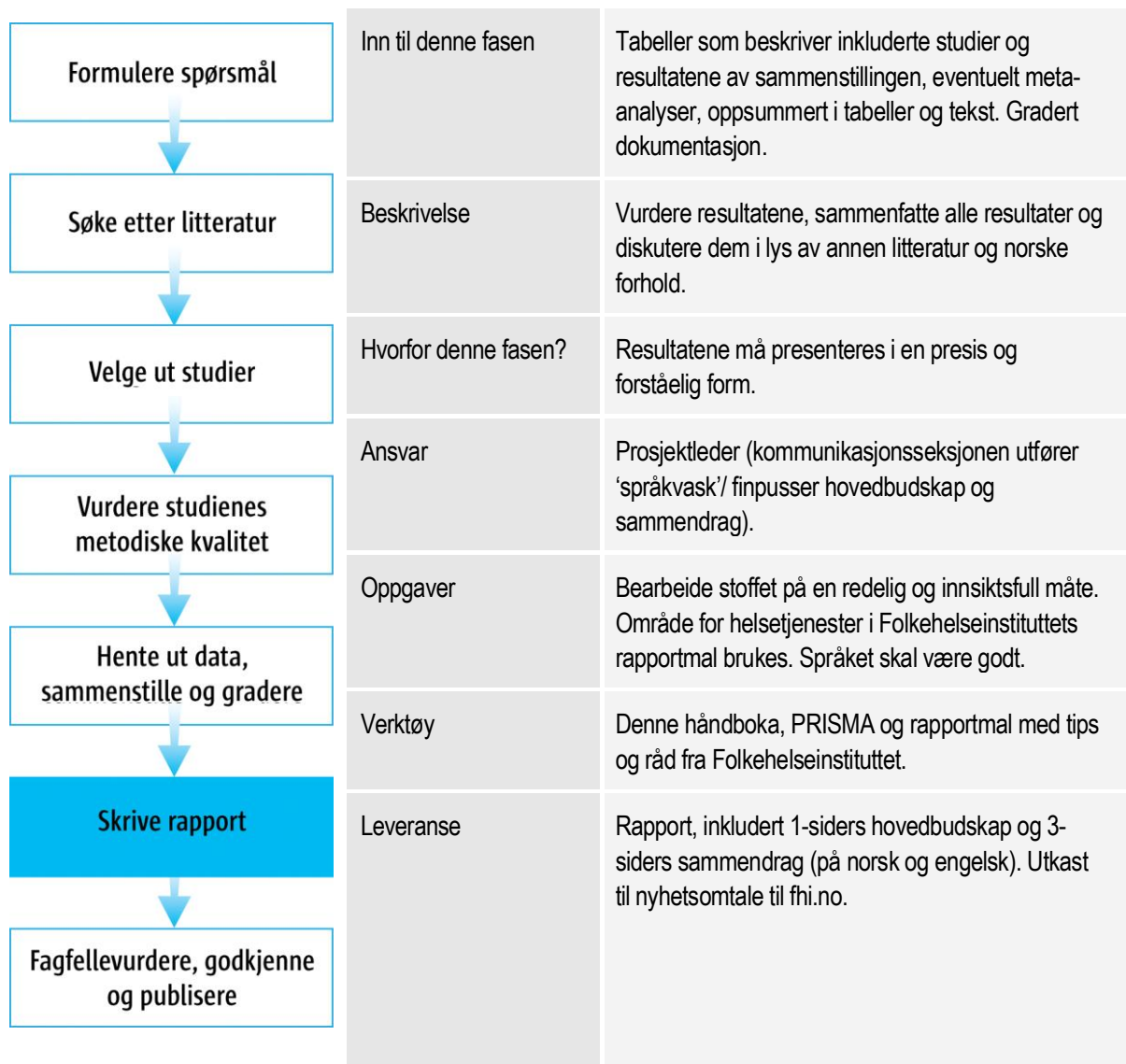
I GRADE håndteres både randomiserte og ikke-randomiserte studiedesign. Der det dreier seg om et effektspørsmål og dokumentasjonen består av randomiserte kontrollerte studier, er dette i utgangspunktet et godt grunnlag for å trekke konklusjoner om effektestimater. Randomiserte studier starter på høy. Dersom det imidlertid er svakheter i måten forsøkene ble utformet og gjennomført på (høy risiko for feil/skjevheter), kan vurderingen bli nedjustert ett trinn, eller ved svært alvorlige svakheter justeres ned to trinn (til lav tillit). Hvis forsøkene i tillegg har relativt få observasjoner (lite data), kan kvaliteten nedgraderes ytterligere (til svært lav tillit).

I et annet tilfelle kan det for eksempel være et effektspørsmål som bare er bedømt ved hjelp av observasjonsstudier (f.eks kohort- eller kasus-kontrollstudier). Da er faren for å feilbedømme effekten større enn ved randomiserte forsøk, og vi starter med lav tillit. Dersom det imidlertid er påvist en sterk sammenheng mellom intervensjon og utfall, kan tilliten til dokumentasjonen oppgraderes ett trinn. Hvis studiene i tillegg tyder på en klar sammenheng mellom dosering av intervensjonen og respons på utfallet, kan

tilliten til dokumentasjonen oppgraderes enda ett trinn og alt i alt vurderes som høy.

GRADE-metoden er beskrevet og illustrert med eksempler i hjelpefilen til programvaren GRADEpro, som er gratis tilgjengelig via <http://grade.pro.org/>, www.gradeworkinggroup.org eller på www.cochrane.org.

8 Skrive oversikten



Vi følger internasjonale standarder for hvordan vi utarbeider og rapporterer systematiske oversikter. For rapportering benytter vi standarden PRISMA, som står for «Preferred Reporting Items for Systematic Reviews and Meta-Analyses». PRISMA er et evidens-basert minimum av elementer som skal rapporteres (beskrives) for en systematisk oversikt.

INNHALDET I RAPPORTENE

Å skrive rapporten eller artikkelen er en viktig del av det å utføre en oversikt. Rapportene lages i ulike formater avhengig av tiden som er til rådighet og hva slags oversikt det dreier seg om. Mens en hurtigoversikt kan være et dokument på noen få sider, vil en systematisk oversikt være mer omfattende. Uansett type oppdrag prøver vi å ta hensyn til behovet for å kommunisere effektivt med oppdragsgiver og helsetjenesten. Det kan være aktuelt å rapportere funnene både i form av en rapport publisert av Folkehelseinstituttet og i artikler publisert i vitenskapelige og internasjonalt tilgjengelige tidsskrift. Det kan være mest hensiktsmessig å skrive på engelsk slik at flest mulig kan forstå og ha nytte av oversiktene.

Tittel

Hvis oversikten omhandler effekt av tiltak bør tittelen skrives på formen "[Tiltak eller sammenlikninger] mot [helseproblem] i eller for [type befolkningsgruppe, sykdom eller problem eller situasjon hvis angitt]".

Hvis alle mulige tiltak skal vurderes i oppsummeringen, kan oversikten ganske enkelt hete "Tiltak mot ..."

Hvis oversikten er avgrenset til en undermengde av mulige tiltak, må tittelen angi denne undermengden. Eksempel: "Primærforebygging av hjerte- og karsykdom, med hovedvekt på medikamentelle tiltak."

Hvis to typer tiltak skal sammenliknes, bør begge tiltak med i tittelen. Eksempel: "Immediate versus delayed treatment for cervical intraepithelial neoplasia."⁸

Unngå å starte med «Effekt av ...» for å unngå at publikasjonslister blir vanskelige å finne fram i for brukerne.

Forfatterskap

En systematisk oversikt av høy kvalitet utarbeides av minst to personer; i langt de fleste tilfellene er det en gruppe som står bak. Det betyr at man tidlig i prosessen må avklare hvem som skal få hvilken anerkjennelse for arbeidet. Kriterier for forfatterskap slik de er utviklet av International Committee of Medical Journal Editors (<http://www.icmje.org>), er:

1. å yte vesentlig til idé og utforming, eller datainnsamling, eller analyse og tolkning av data OG

⁸ Se også Cochrane Handbook: <http://www.cochrane-handbook.org/>

2. ha ført rapportutkastet i pennen eller kritisk revidert viktig intellektuelt innhold OG
3. ha gitt endelig godkjenning av versjonen som skal publiseres OG
4. samtykke til å stå ansvarlig for alle deler av arbeidet for å sikre at spørsmål knyttet til presisjon eller integritet til enhver del av arbeidet er hensiktsmessig undersøkt og løst

For å kvalifisere som forfatter av en oversikt må man altså ha bidratt på alle fire punkter. I området bruker vi disse kriteriene, og det har flere konsekvenser.

For det første kan vi etter mønster av det som nå er vanlig i mange tidsskrifter beskrive hvem av forfatterne som har gjort hva i prosessen. Det gjør vi i forordet.

For det andre bør alle som deltar med vesentlige bidrag i punkt 1, inviteres av prosjektleder til også å delta i de neste tre punktene slik at de kan tas med som forfattere. Eksempelvis bør bibliotekarer som inngår i prosjektet, inviteres til å kommentere og gi innspill på manuskriptets intellektuelle innhold (punkt 2), og til å godkjenne den endelige versjonen (punkt 3). Bibliotekaren må så – i likhet med de resterende forfatterne – være innforstått med forpliktelsen som ligger i punkt 4. Eventuelle faglige uenigheter kan ikke løses ved at medforfattere bare tar ansvar for deler av manuskriptet.

For det tredje kan de som ikke fyller alle forfatterkriteriene, men likevel har bidratt, i stedet takkes i forordet. Eksempel på bidrag som ikke oppfyller forfatterkriteriene, men som kan takkes for i forordet:

- generell oppfølging (supervisjon) eller administrativ støtte til prosjektgruppa
- fagfelleevaluering
- skriveassistanse, teknisk redigering, nøyelesing, språkvask og korrekturlesing

ICMJE-anbefalingene omtaler ikke rekkefølgen på forfatterne. Vi følger prinsippet om at arbeidsinnsatsen avgjør. Prosjektleder vil vanligvis være førsteforfatter, og medforfatterne følger i nedstigende rekke etter arbeidet som er nedlagt. Hvis tvil, diskuterer forfatterne seg til enighet som kollegium.

Endelig utforming av rapporten

For å få en felles form på oversiktene våre, har vi utarbeidet en felles mal for rapportene. Den sørger for et visuelt enhetlig uttrykk i bruk av fonter, overskrifter og avsnitt som gjør det tydelig at dette er et produkt fra område for helsetjenesten i Folkehelseinstituttet. Malen inneholder veiledning til hvert kapittel og eksempeltekster til hjelp i skriveprosessen.

Kommunikasjonsseksjonen har i tillegg utarbeidet skrivetips for at språket i rapportene skal bli så godt som mulig. Veiledningen som gis forfattere i Tidsskrift for Den norske legeforening, kan også være en kilde til gode råd (22).

Skrivetips

Skriv kort, unngå fremmedord

- Tenk på hvem du skriver for
 - Velg ord og uttrykk som leseren kan forstå
 - Forklar faguttrykk når du må bruke det
- Vis hvem som handler
- Bruk verb til å beskrive de viktigste handlingene
- Sørg for at teksten henger logisk sammen
 - Plasser kjent informasjon først i setningen/avsnittet
 - Plasser ny/ukjent informasjon i slutten av setningen/avsnittet
- Det er ingen skam å sette punktum
- Har du flere ting på hjertet, så si én ting om gangen
- Vær konkret – mener du en spade, så skriv en spade

Vancouver-systemet. Dette gjelder generelt for alle referanser:

- referansene skal være nummerert og ordnet i den rekkefølgen de opptrer i teksten.
- bruk små bokstaver i tittelen (også på engelsk), bortsett fra ved første ord i ny setning og ved egennavn.
- hvis mange forfattere, oppgi navnet på de seks første etterfulgt av et al.

Du finner utfyllende informasjon om hvordan ulike dokumenttyper bør refereres i rapportmalen. Mer informasjon om bruk av EndNote finner du på intranettet > Bibliotek tjenester.

Husk å kvalitetssjekke referansene før publisering.

Du kan også lese mer på National Library of Medicine (24).

9 Fagfelleverdere, godkjenne og publisere

Formulere spørsmål	Inn til denne fasen	Utkast til rapport.
↓	Beskrivelse	Alle oversikter gjennomgår fagfellevurdering. Hvor omfattende denne er avhenger av oversiktstype.
Søke etter litteratur	Hvorfor denne fasen?	Å sikre at rapporten er metodisk god og relevant for norsk praksis og at vårt arbeid gjøres kjent.
↓	Ansvar	Prosjektleder.
Velge ut studier	Oppgaver	Sørge for at rapporten fagfellevurderes og sendes til godkjenning.
↓	Verktøy	CRD Guidance. Områdets rapportmal, skjema for fagfellevurdering, Områdets informasjonsstrategi, manuskriptveiledning i enkelttidsskrifter og i de aktuelle internasjonale samarbeidene (Cochrane m.fl.).
Vurdere studienes metodiske kvalitet	Leveranse	Kommentarer til rapporten. Vedtak om godkjenning. Elektronisk versjon for nettpublisering.
↓		
Hente ut data, sammenstille og gradere		
↓		
Skrive rapport		
↓		
Fagfelleverdere, godkjenne og publisere		

FAGFELLEVRDERING

Fagfellevurdering er en måte å bedre kvaliteten av forskningsprodukter på, som typisk brukes når man publiserer forskning i form av en tidsskriftartikkel. Område for helsetjenster i Folkehelseinstituttet ønsker i størst mulig grad å innhente eksterne vurderinger fra temaeksperter også av våre oversikter, ikke minst for å gjøre beskrivelsen av kliniske aspekter så god som mulig. Vi benytter også intern fagfellevurdering, særlig for å vurdere metode i samsvar med kriteriene utarbeidet av PRISMA-gruppen, og språk.

PUBLISERE OVERSIKTENE

Alle våre rapporter legges ut på våre nettsider, og vi tar gjerne imot kommentarer til ferdigstilte oppsummeringer. Dersom det blir påvist feil, vil rapporten bli endret.

I tillegg publiseres et engelsk sammendrag av våre rapporter i CRD-databasen som et INATHA brief.

I de tilfeller rapportene også skrives om og publiseres som artikler, så oppmuntres det til publisering i open access-tidsskrifter.

10 Diagnostikk

NB! Dette kapittelet vil bli ferdigstilt på et senere tidspunkt.

Kunnskapssenteret bruker QUADAS-2 for vurdering av studier som rapporterer diagnostiske utfallsmål.

For studier som rapporterer helseeffekter av diagnostiske tester, bruker vi risiko for systematiske feil (bias). Kvaliteten på dokumentasjonen kan graderes med GRADE. Se kapittel 8, side 50-53.

Liv Merete spør Kjetil GB om han vil skrive kapittelet om diagnostikk

Signe spør Atle som spør Claire og Simon om kapittel om Kvalitative oppsummeringer

Hvem skal skrive noe om brukermedvirkning? Skal det også inn her? eller i Prosjektleders admin håndbok?

Skal vi gi en 'anbefaling' om at det kan være nyttig å bruke DECIDE som hjelpemiddel videre i prosessen til å ta beslutninger?

11 Kvalitative oppsummeringer

NB! Dette kapitlet vil bli ferdigstilt på et senere tidspunkt.

12 Metodevurderinger

INNLEDNING

NB! Dette kapittelet vil bli ferdigstilt på et senere tidspunkt.

En metodevurdering inneholder i tillegg til en systematisk oversikt om effekt av tiltak også analyser av konsekvenser ved å innføre en ny metode eller endre praksis. Dette kan være analyser av økonomiske, organisatoriske, etiske og juridiske konsekvenser for pasient og samfunn.

Den delen av en metodevurdering som er en systematisk oversikt, følger fremgangsmåten beskrevet i kapitlene 4 til 8. Avhengig av hvilket tiltak eller hvilken problemstilling som vurderes, vil vi belyse ulike konsekvenser av å innføre tiltaket. Oftest gjør vi en helseøkonomisk evaluering, men etiske konsekvenser blir også jevnlig belyst. Hvordan vi utfører dette, er beskrevet i dette kapittelet.

Etter at Nasjonalt system for innføring av nye metoder i spesialisthelsetjenesten ble innført, er metodevurdering blitt en samlebetegnelse for flere produkter:

- **Fullstendig metodevurdering:** mer omfattende systematisk vurdering av nye eller etablerte metoder der både effekt, sikkerhet og kostnadseffektivitet undersøkes.
- **Hurtig metodevurdering:** mer avgrenset enn en fullstendig metodevurdering; baseres ofte på dokumentasjon sammenfattet av produsenten og utarbeides vanligvis innen 180 dager fra mottak av denne dokumentasjonen.
- **Mini-metodevurdering:** en forenklet metodevurdering, som benyttes lokalt i helseforetakene som beslutningsstøtteverktøy. Vurderingen utføres lokalt ved sykehusene og involverer forslagsstillerene (vanligvis klinikere), fagfeller og beslutningstakere, som fyller ut et skjema for mini-metodevurdering, utarbeidet ved Kunnskapssenteret.

I en økonomisk evaluering blir kostnader og effekter av medisinske tiltak (metoder, virkemidler, programmer) sammenliknet, enten for å innføre nye medisinske helsetiltak eller utvide eller innskrenke eksisterende tilbud. Hensikten med en økonomisk analyse er å legge premisser for en mest mulig effektiv utnytting av ressurser. Med effektiv ressursutnytting mener vi:

- At ressursene brukes slik at vi oppnår mest mulig (ut fra målene) for en gitt mengde ressurser
- Eller at vi oppfyller et gitt mål med minst mulig bruk av ressurser

Det kan være hensiktsmessig å skille mellom effektivitet på to nivåer: Nivået der prioriteringene gjøres (samfunnsnivået) og nivået der politikken iverksettes (driftsnivået).

Både politikere som prioriterer på makronivå og beslutningstakere som iverksetter politikken har følgelig nytte av økonomiske beregninger. Kunnskapssenteret gjør økonomiske evalueringer både sett fra helsesektoren og samfunnets perspektiv. Vi beregner også kostnadene for aktører som iverksetter eller berøres av politikken (institusjoner, pasienter osv.).

Generelt om økonomisk evaluering

For prioriteringsformål ønsker vi å vite hvor mye større eller mindre effekten av det ene helsetiltaket er enn av det andre, og hvor mye mer eller mindre helsetiltaket koster enn et annet tiltak i helsetjenesten (se tabellen nedenfor).

Endring i kostnader og effekt/nytte når eksisterende helsetiltak erstattes med nytt

	Effekten av nytt helsetiltak i forhold til eksisterende	
Kostnaden	MINDRE	STØRRE
LAVERE	A. Beregn kostnad-effekt-brøk	B. Velg det nye tiltaket (dominerende strategi)
HØYERE	C. Behold det gamle tiltaket (dominerende strategi)	D. Beregn kostnad-effekt-brøk

Valget er greit i situasjon B og C. I de to andre tilfellene er det ikke umiddelbart klart om det nye tiltaket bør foretrekkes: I situasjon A er både kostnader og

effekt lavere enn for eksisterende tiltak, mens i D er begge høyere. For slike situasjoner beregner vi forholdet mellom kostnad og effekt. Det uttrykker hvor mye ekstra et tiltak koster for å oppnå én ekstra effektenhet (et vunnet, eventuelt kvalitetsjustert, leveår) i forhold til et annet tiltak.

For øvrig kan det tenkes at samfunnet ønsker å prioritere ikke bare ut fra kostnader i forhold til effekt eller nytte av behandling, men for eksempel også etter hvilke tiltak som gir størst effekt alene. Det kan også tenkes at samfunnet ønsker å prioritere etter alvorlighetsgraden av en sykdom.

Det er en utfordring å basere alle økonomiske analyser på pålitelige estimater av tiltakets effekt. Det hjelper lite å ha gode data for økonomi hvis effektestimatet ikke er korrekt. Det er derfor viktig å basere effektvurderingen i størst mulig grad på gode systematiske oversikter med kvantitativ oppsummering (meta-analyser) av randomiserte kontrollerte forsøk.

Metoder i økonomiske analyser

For å beregne forholdet mellom kostnad og effekt er ulike metoder aktuelle (12).

Kostnadseffektivitetsanalyse (cost-effectiveness analysis, CEA)

Hvis effektmålet er klinisk effekt alene, og effektmålet er likt for de to helsetiltakene, er kostnadseffektivitetsanalyse tilstrekkelig. I denne modellen beregner vi kostnad per klinisk effektenhet, for eksempel vunnet leveår, redusert sykkelighet, symptomreduksjon eller unngåtte hendelser (for eksempel hoftebrudd). Her tas det hensyn til to ulike typer kostnader:

- Kostnader for helsesektoren (direkte kostnader)
- Eventuelle kostnader som faller på pasienten, deres pårørende eller andre samfunnssektorer (indirekte kostnader)

Kostnadene for helsesektoren må alltid tas med. Når perspektivet for analysen er samfunnsøkonomisk, ønsker vi også i størst mulig grad å trekke inn alle kostnader for samfunnet (både direkte og indirekte kostnader).

QALY-analyse (kostnad–livskvalitetsanalyse, cost-utility analysis, CUA)

Hvis målet derimot er knyttet til livskvalitet eller verdsetting av den kliniske effekten (og dette er målbart), og/eller de to metodene har forskjellig utfallsmål (ulike mål på klinisk effekt), benyttes QALY-analyse eller kostnad–livskvalitetsanalyse. Her er effektmålet såkalte kvalitetsjusterte leveår (Quality Adjusted Life Years - QALYs), som fanger opp effekter på både livskvalitet og leveår. Livskvalitetsforbedringer ved bruk av et helsetiltak krever en form for verdsetting av helseforbedringene. Denne verdsettingen kan gjøres av pasienten eller av samfunnet. Størrelsen på effekten vil kunne variere mye

etter hvilket tiltak som er brukt i verdsettingen, og det kan derfor være klokt å konferere med en erfaren helseøkonom når man skal tolke slike studier.

Som kostnader regner vi her de direkte kostnadene for helsesektoren og eventuelle indirekte kostnader (som i en kostnadseffektivitetsanalyse, CEA). Som i en CEA ønsker vi å inkludere alle kostnader for samfunnet når perspektivet er det samfunnsøkonomiske.

Kostnad–nytte-analyse (cost-benefit analysis, CBA)

I kostnad–nytte-analysen beregnes samfunnets kostnader og samfunnets nytte av et virkemiddel eller tiltak. Med samfunnets kostnader menes kostnadene for helsesektoren og kostnadene for pasient, pårørende og samfunnet for øvrig (indirekte kostnader). Med nytte menes samfunnets betalingsvillighet (i kroner) for å ha et helsetjenestetilbud tilgjengelig. Nytte i denne forstand er vanskelig målbart og lite brukt for vurdering av tiltak i helsesektoren.

Betydningen av klinisk effekt

All økonomisk analyse vi gjør, tar utgangspunkt i den dokumenterte effekten. Hvilken tilnærming vi velger, avhenger også av hvem oppdragsgiver er og vedkommendes ønsker.

Ingen dokumentert effekt

Når oversikten konkluderer med at helsetiltaket er uten dokumentert effekt eller muligens skadelig, er det lite behov for en økonomisk evaluering fordi vi uansett ikke bør prioritere tiltaket.

Dersom det aktuelle virkemidlet eller tiltaket allerede er i bruk, kan det være aktuelt å beregne de driftsøkonomiske kostnadene knyttet til bruken av helsetiltaket (f.eks. i sykehus). Et estimat av disse kostnadene vil kunne være verdifullt for bl.a. sykehusledelsen for å kunne drive sykehus effektivt i økonomisk forstand. Når et tiltak allerede er i bruk, kan det også være relevant å beregne de samfunnsøkonomiske kostnadene forbundet med eksisterende tilbud, eller beregne kostnader knyttet til den aktuelle sykdommen.

Usikker effekt

I tilfeller der den kliniske effekten er usikker, kan det være aktuelt å gjøre en økonomisk evaluering som illustrerer konsekvensene av ulike, usikre effektestimater. Vi gjør med andre ord en **sensitivitetsanalyse** for å se hvordan resultatet (kostnad–effekt-forholdet) påvirkes av variasjoner eller endringer i premissene som legges til grunn for analysen (i denne sammenhengen antatt effekt). Det kan også være aktuelt å beregne kostnader som i punktet over.

Dokumentert effekt

I tilfeller med dokumentert klinisk effekt, gjør vi en **økonomisk evaluering**. Det kan også være aktuelt å beregne de samfunnsøkonomiske kostnadene og de driftsøkonomiske kostnadene, og eventuelt beregne kostnader knyttet til den aktuelle sykdommen. Det kan nemlig være slik at en metode er lønnsom for samfunnet og fortjener prioritet, men at man ikke har ressurser verken på samfunnsnivå eller bedriftsnivå (for eksempel sykehusnivå) til å iverksette metoden for alle syke, i alle fall ikke på kort sikt.

Omfang og tilnærming

Hvor mye tid og ressurser vi setter av til å gjøre en økonomisk evaluering, avhenger blant annet av hvor godt dokumentert effekten er og av hva som har vært gjort tidligere av denne typen analyse på det aktuelle feltet.

Evaluerings basert på publiserte økonomiske analyser

En del helsetiltak har tiltrukket seg mye oppmerksomhet, og det er publisert mange økonomiske evalueringer. I slike tilfeller kan det være aktuelt å søke etter litteratur og kritisk vurdere de aktuelle studiene. Det er utviklet sjekklister for slike studier som et redskap til bruk i gjennomgang av helseøkonomisk litteratur (se vedlegg 2). Vi velger ut analyser med høy intern validitet som passer for den aktuelle settingen og gjør en økonomisk vurdering basert på dette materialet. Et problem er at kostnader, risiko og forekomst for ulike hendelser varierer mellom land, og at helsetjenesten er organisert forskjellig. Det vil således være større usikkerhet knyttet til internasjonal dokumentasjon på helseøkonomiske data enn for kliniske effektdata.

Fullstendig evaluering

Når en systematisk oversikt konkluderer med relativt presise effektestimater basert på rimelig gode studier, kan det være aktuelt å lage en fullstendig økonomisk analyse. Basert på tilgjengelige data for kostnader, nytte og risiko for ulike hendelser, lager vi en modell som brukes til å beregne kostnad–effekt-forhold. Fordelen med slike modeller er at vi kan gjøre sensitivitetsanalyser til å kartlegge hvordan resultatet påvirkes av endringer i forutsetningene vi gjør om variablene i modellen. Ved å endre på premissene eller estimatene vi legger til grunn for modellen, får vi nye kostnad–effekt-forhold. Modellen er dermed også egnet til bruk i andre land. Sensitivitetsanalysen kan også gi informasjon om et bestemt helsetiltak er lønnsomt å ta i bruk for noen pasientgrupper og ikke for andre. Andre ganger kan store variasjoner i resultatet som følge av endringer i en variabel være et uttrykk for at modellen bør forbedres ved at man søker et mer presist estimat for vedkommende variabel. Ved utvikling av egen modell er det viktig å ha i tankene alt som står i vår egen sjekkliste for slike studier (se vedlegg 2).

Bruk av en utenlandsk modell med norsk tilpasning

Dersom det finnes gode utenlandske modeller for økonomisk evaluering av medisinske metoder, kan det være aktuelt å «låne» en utenlandsk modell og lage beregninger basert på norsk medisinsk praksis og norske kostnadstall.

Hvordan vi lager en helseøkonomisk evaluering

Helseøkonomiske evalueringer lages i forbindelse med metodevurderinger og parallelt med oppsummeringen. For at resultatet skal bli best mulig, er det viktig at helseøkonomisk kompetanse involveres så tidlig som mulig. Vi har derfor her skissert hvordan dette gjøres, og knyttet dette opp til de forskjellige fasene.

Fase 1

I alle metodevurderinger bør man vurdere å ha med noe om helseøkonomi. Hvis man mener at dette ikke er relevant, diskuteres det med en helseøkonom, og man bør gjøre rede for beslutningen i den endelige rapporten. For at det helseøkonomiske bidraget skal være så relevant som mulig, bør helseøkonomisk kompetanse involveres allerede ved bestilling og skriving av prosjektplan.

Fase 2

De som skal jobbe med det helseøkonomiske bidraget, involveres i utarbeidingen av søk etter helseøkonomisk litteratur hvis en oppsummering av dette skal være en del av rapporten. Det er viktig at man blir enige om de forskjellige aspektene ved søket (søkefiltre, databaser, strategier osv.).

Fase 3

Inklusjon/eksklusjon av helseøkonomiske studier gjøres basert på de samme kriterier som for andre studier; populasjon, intervensjon, sammenlikning, utfall og studiedesign. Hva som er relevant, diskuteres før artiklene foreligger.

Fase 4

Når man har konkludert med hvilke studier som skal inkluderes, bør skissen til en eventuell helseøkonomisk modell ferdigstilles. Valg av type tilnærming er beskrevet tidligere.

Fase 5

Hvis man gjør en oppsummering av helseøkonomisk litteratur, hentes det ut data basert på sjekklisten i vedlegg 2. Hvis det gjøres en egen analyse, henter man inn de tallene man trenger. Det gjelder både eventuelle effektdata fra oppsummeringen og andre data, som kostnader og overgangssannsynligheter. Deretter gjøres analysene.

Fase 6

Rapporten skrives og behandles på samme måte som andre rapporter.

Andre aspekter

En god helseøkonomisk evaluering bør gi svar på det spørsmålet som er stilt. Den bør også drøfte andre aspekter enn dem som kommer til tallmessig uttrykk i resultatet. Reiser analysen etiske spørsmål med hensyn til hvilke pasientgrupper som prioriteres opp eller ned? Er det andre faktorer enn kostnader og helsenytt som kan tenkes å påvirke den beslutningen analysen skal gi grunnlag for?

ETISKE VURDERINGER DET ER FORESLÅTT Å BYTTE UT MORAL OG MORALSKE MED ETISK OG ETISKE, OK?

Hvorfor er etikk en del av metodevurderinger?

Helt fra starten har etikk vært definert som en del av metodevurderinger, og det finnes mange grunner til at etikk har en selvfølgelig plass der (13). Det viktigste er at alle tiltak for diagnostisering og behandling har et moralsk mål: Å gi mennesker *et bedre liv*. Dessuten er metodevurderinger en *vurdering* som baserer seg på moralske så vel som faglige normer og verdier. Ofte henger disse nøye sammen, for eksempel i valg av endepunkter. Nye metoder utfordrer også moralske normer og verdier, og hvis de er kostbare, kan innføringen av dem få implikasjoner for fordelingen av ressurser i helsevesenet.

Beslutninger om å bruke ulike tiltak i helsevesenet bygger derfor ikke bare på kunnskap om sikkerhet, effektivitet og kostnadseffektivitet, men også på vurderinger av moralske aspekter.

Hva er målet med etikk i metodevurderinger?

Å inkludere etiske vurderinger kan ha flere mål:

1. Synliggjøre og informere om moralske aspekter ved innføringen av en ny metode/intervensjon.
2. Drøfte etiske argumenter som kommer fram i vurdering av en metode
3. Gi råd om innføring av et tiltak ut fra de moralske aspektene og vurderingene.

Kunnskapssenteret har som mål å innhente, analysere og formidle kunnskap som inngår i beslutninger på alle nivåer i helsetjenesten, men gir ikke konkrete råd. Etikken vil derfor konsentrere seg om punkt 1 og 2.

Hvordan skal etikk være en del av metodevurderinger?

Det er utviklet en rekke metoder for håndtering av moralske spørsmål i metodevurderinger (14;15). Kunnskapssenteret anvender en metodikk som er spesielt utviklet for å synliggjøre og drøfte moralske aspekter ved helsefaglige tiltak (14). Metoden utgjør en byggestein i en europeisk kjernemodell for metodevurderinger (16;17) og er samstemt med denne.

Hvordan går vi fram?

Her følger kort en beskrivelse av en praktisk framgangsmåte for å belyse og drøfte moralske aspekter ved en teknologi eller metode. Kjernen i metoden er en rekke spørsmål, som er listet i vedlegg 2.

Praktisk framgangsmåte:

1. Beskrivelse og analyser av helsetiltaket som skal vurderes med hensyn på spesifikke moralske utfordringer

Svarene kan få følger for valg av endepunkter og det øvrige arbeidet med vurdering av helsetiltak.

2. Identifisering av berørte parter og interessenter

Vurdering av hvordan de berørte partene blir hørt og tatt hensyn til i metodevurderingen og vurdering av omfanget av litteratursøkingen.

3. Valg av spørsmål

Vurdering av hvilke spørsmål som er relevante.

4. Litteratursøking

med tanke på å besvare spørsmålene som er valgt ut i punkt 3.

5. Sammenfatning av prosessen

Med unntak av litteratursøking (punkt 4) gjennomfører prosjektlederen alle trinn i nært samarbeid med utredningsgruppa som gjennomfører oppsummeringen, såfremt en slik gruppe er oppnevnt. Prosessen kan være syklisk. Man går tilbake og vurderer for eksempel endepunktvalg på nytt dersom man under drøftingen av spørsmålene oppdager aspekter ved helsetiltaket som man ikke har fanget opp tidligere. For nærmere beskrivelse av metoden og dens bakgrunn, se referansene 14 og 18-21.

Hvem kan belyse og drøfte moralske aspekter?

Selv om det er en fordel at man er skolert i etikk, er metoden utviklet for at den skal kunne brukes også uten formell bakgrunn i etikk.

JURIDISKE BETRAKTNINGER, ORGANISATORISKE BETRAKTNINGER, ANDRE BETRAKTNINGER?

Vurdering av juridiske konsekvenser kan være så enkelt som å vurdere om metoden har nødvendig godkjenning til bruk i rutine, eller om innføring av metoden vil kreve endringer i det eksisterende lovverket.

Dersom det skal gjøres mer omfattende vurderinger, vil vi søke ekstern ekspertise på dette feltet.

Vurdering av organsatoriske konsekvenser er særlig viktig dersom metoden endrer tidligere rutiner, krever investeringer i nye fasiliteter, forutsetter oppbygging av kompetanse, berører funksjonsdeling eller påvirker samhandling med andre, for eksempel mellom spesialist og primærhelsetjenesten.

Ofte vil det være nødvendig for oss å samarbeide med eksterne fagpersoner for å få dette grundig vurdert.

Litteratur

1. Systematic Reviews: CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination, University of York; 2009. Tilgjengelig fra: [25.05.2011]
2. Higgins J, Green S, editors. Cochrane handbook for systematic reviews of interventions. Version 5.1.0. [Updated March 2011]. The Cochrane Collaboration; 2011. Tilgjengelig fra: <http://www.cochrane-handbook.org/> [25.05.2011]
3. Bjørndal A, Flottorp S, Klovning A. Kunnskapshåndtering i medisin og helsefag. Oslo: Gyldendal Akademisk; 2007.
4. Guyatt G, Rennie D, editors. Users' guide to the medical literature. A manual for Evidence-Based Clinical Practice. Chicago: American Medical Association; 2002.
5. EUnetHTA. HTA Core Model Handbook. EUnetHTA Work Package 4 (WP4). Tilgjengelig fra: <https://fio.stakes.fi/htacore/handbook.html> [25.05.2011]
6. Jamtvedt G, Hagen K, Bjørndal A. Kunnskapsbasert fysioterapi. Oslo: Gyldendal akademisk; 2003.
7. Bjørndal A, Hofoss D. Statistikk for helse- og sosialfagene. Oslo: Gyldendal akademisk; 2004.
8. Bjørndal A. Tenke- og arbeidsmåter i medisinen – en innføring i den medisinske logikk. Oslo: Gyldendal akademisk; 2003.
9. Shadish W, Cook T, Campbell D. Experimental and quasi-experimental designs for Generalized Causal Inference. Boston: Houghton Mifflin Company; 2002.
10. Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;336(7644):601-5.
11. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336(7650):924-6.
12. Drummond M, Sculpher MJ, Torrance G, O'Brien BJ, Stoddart G. Methods for the economic evaluation of health care programmes, third edition. Oxford: Oxford University Press; 2005.

13. Hofmann BM. Why ethics should be part of health technology assessment. *Int J Technol Assess Health Care* 2008;24(4):423-9.
14. Hofmann B. Etikk i vurdering av helsetiltak. Utvikling av en metode for å synliggjøre etiske utfordringer ved vurdering av helsetiltak. Oslo: Kunnskapssenteret i Folkehelseinstituttet; 2008. Tilgjengelig fra: <http://www.kunnskapssenteret.no/publikasjoner/etikk-i-vurdering-av-helsetiltak.utvikling-av-en-metode-for-a-synliggjore-etiske-utfordringer-ved-vurdering-av-> [25.11.2015]
15. Saarni SI, Hofmann B, Lampe K, Luhmann D, Makela M, Velasco-Garrido M, et al. Ethical analysis to improve decision-making on health technologies. *Bull World Health Organ* 2008;86(8):617-23.
16. Lampe K, Mäkelä M, editors. HTA Core Model for Medical and Surgical Interventions 1.0 R. Work Package 4. Helsinki: European Network for Health Technology Assessment (EUnetHTA); 2008. Tilgjengelig fra: http://www.eunethta.eu/en/Public/Work_Packages/EUnetHTA-Project-2006-08/WP_4/Activities/ [25.05.2011]
17. Lampe K, Mäkelä M, editors. HTA Core Model for Diagnostic Technologies 1.0 R. Work package 4. Helsinki: European Network for Health Technology Assessment (EUnetHTA); 2008. Tilgjengelig fra: http://www.eunethta.eu/en/Public/Work_Packages/EUnetHTA-Project-2006-08/WP_4/Activities/ [25.05.2011]
18. Hofmann B. On value-judgements and ethics in health technology assessment. *Poiesis Prax* 2005;3(4):277-95.
19. Hofmann B. Toward a procedure for integrating moral issues in health technology assessment. *Int J Technol Assess Health Care* 2005;21(3):312-8.
20. Hofmann B. When means become ends: technology producing values. *Seminar net - Media, technology and lifelong learning* 2006;2(2). Tilgjengelig fra: <http://seminar.net/index.php/volume-2-issue-2-2006-previousissuesmeny-114/66-when-means-become-ends-technology-producing-values> [25.05.2011]
21. Hofmann B. Is there a technological imperative in health care? *Int J Technol Assess Health Care* 2002;18(3):675-89.
22. Tidsskrift for den norske legeforening. Tidsskriftets veiledning for forfattere og fagvurderere. Tilgjengelig fra: http://m.tidsskriftet.no/index.php?nota_id=480 [25.05.2011]
23. Helsedirektoratets bibliotek. Veiledning—Hvordan lage referanseliste. [Oppdatert 15.02.2011]. Tilgjengelig fra: <http://www.helsedirektoratet.no/fagnytt/bibliotek> [25.05.2011]
24. National Library of Medicine. International Committee of Medical Journal Editors (ICMJE) Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals: Sample References. [Oppdatert 04.11.2015]. Tilgjengelig fra: http://www.nlm.nih.gov/bsd/uniform_requirements.html [25.11.2015]

