

## Statusartikel

# Kritisk læsning af systematiske oversigtsartikler

David Ruben Teindl Laursen<sup>1, 2</sup>, Stig Brorson<sup>3</sup>, Andreas Lundh<sup>1, 2, 4</sup>, Ann Merete Møller<sup>5</sup>, Camilla Hansen Nejtgaard<sup>1, 2</sup>, Jacob Rosenberg<sup>6</sup> & Asbjørn Hróbjartsson<sup>1, 2</sup>

1) Cochrane Danmark & Center for Evidensbaseret Medicin Odense (CEBMO), Klinisk Institut, Syddansk Universitet, 2) Open Patient data Explorative Network (OPEN), Odense Universitetshospital, 3) Center for Evidensbaseret Ortopædkirurgi (CEBO), Sjællands Universitetshospital, Køge, 4) Afdeling for Lunge- og Infektionssygdomme, Københavns Universitetshospital – Bispebjerg og Frederiksberg Hospital, 5) Cochrane Anaesthesia, Critical and Emergency Care Group, Bedøvelse, Operation og Intensiv Behandling Herlev, Københavns Universitetshospital – Herlev Hospital, 6) Cochrane Colorectal Group, Afdeling for Mave-, Tarm- og Leversygdomme, Københavns Universitetshospital – Herlev Hospital

Ugeskr Læger 2024;186:V09230616. doi: 10.61409/V09230616

### HOVEDBUDSKABER

- En systematisk oversigt giver overblik over primærstudier, som undersøger samme forskningsspørgsmål, f.eks. effekten af en given behandling.
- En metaanalyse kan indgå og er et vægtet gennemsnit af de enkelte studieresultater.
- Oversigten bør tage højde for studiers risiko for bias og heterogenitet.

Denne artikel giver en indføring i kritisk læsning af en systematisk oversigtsartikel. Kernebegreber er forklaret i **Tabel 1**, som der henvises til undervejs i teksten. Artiklen tager udgangspunkt i følgende patientcase og efterfølgende relevante eksempel på en systematisk oversigt.

**TABEL 1** Alfabetisk liste over kernebegreber i systematiske oversigtsartikler.

Begreb	Definition/forklaring
AMSTAR 2	Et standardiseret værktøj til at hjælpe med kritisk at vurdere systematiske oversigter [1]
Bloksøgning	En søgestrategi, som opbygges af blokke, der hver dækker et aspekt af forskningsspørgsmålet, f.eks. population eller intervention, og indeholder forskellige synonymmer og emneord for dette aspekt, f.eks. surgery eller operation
Cochraneoversigt	En systematisk oversigt, som er udarbejdet under Cochranesamarbejdet med særligt strenge krav til planlægning, metoder, gennemførelse og udgivelse
Fixed effect-metaanalyse	En metaanalyse, hvor det antages, at al variation kun skyldes måleusikkerhed og ikke heterogenitet, dvs. studierne antages at have den samme underliggende effekt Kan være relevant ved ensartede studier udført under samme protokol
Funnel plot	En figur, hvor effekten fra hvert studie (uden konfidensinterval) plottes på x-aksen med deres måleusikkerhed op ad y-aksen Da måleusikkerheden afhænger af størrelsen på studierne, vil punkterne ofte danne en tragt (eng.: funnel) Hvis f.eks. små studier, som ikke viser gavnlige effekt, ser ud til at »mangle«, vil plottet have et hul og blive asymmetrisk, hvilket ofte fejlagtigt fortolkes som et klart tegn på publikationsbias Asymmetrien kan dog lige så vel skyldes andre forskelle mellem store og små studier, f.eks. metodologiske og kliniske forskelle
GRADE	Et standardiseret værktøj til at hjælpe med at vurdere tiltroen til evidensen (se »Tiltro til evidensen«) [2]
Grå litteratur	Litteratur, der ikke er publiceret i almindelige fagfællede tidsskrifter, men f.eks. som conferenceabstracts, i afhandlinger, i forsøgsregistre, på lægemiddelvirksomheders hjemmesider, på lægemiddelstyrelsens hjemmesider og som preprints
Hurtigoversigt (eng.: rapid review)	En variant af systematisk oversigt, hvor forskerne prioriterer hurtig færdiggørelse, hvis et hurtigt svar er nødvendigt, f.eks. ved en pandemi, på bekostning af f.eks. grundigere søgning eller dataindsamling af 2 uafhængige personer
Klinisk relevant forskel	En forskel i effekt, som er $\geq$ den mindste klinisk relevante forskel, og som bør have konsekvens for valg af behandling Det kunne f.eks. være $\geq 1,5$ points forskel på en visuel analog skala på 0-10 for smerteintensitet Den mindste klinisk relevante forskel kan bestemmes på forskellige måder, f.eks. ved at sammenligne en forbedring på skalaen med, hvornår patienterne oplever en overordnet bedring af funktion
Kortlægningsoversigt (eng.: scoping review)	En variant af systematisk oversigt, hvis formål er deskriptiv kortlægning af tilgængelige studier og af videnskuller og ikke f.eks. et samlet svar på effekten af en behandling Det kan f.eks. være: Hvordan udformes forsøg, der undersøger effekt af skulderoperationer?
Metaanalyse	En statistisk analyse til beregning af et vægtet gennemsnit, hvor resultaterne fra de enkelte studier indgår med en vægt afhængig af deres statistiske præcision Resultatet af en metaanalyse vil ofte blive præsenteret i et forest plot En metaanalyse kan være en random effects-metaanalyse eller fixed effect-metaanalyse
Metaanalyse med individuelle patientdata	En variant af metaanalyse, hvor studierne indgår med deres komplette datasæt på patientniveau og ikke kun summariske data Dette muliggør en standardisering af analysen og mere detaljerede analyser af årsager til heterogenitet
Metaregression	En analyse, hvor metaanalysen inddrager en (ofte kontinuert) faktor, f.eks. varierende dosis af medicin, til at forsøge at forklare heterogenitet En eventuel ændring i effekt afhængig af den undersøgte faktor kaldes også effektmodifikation

**TABEL 1 FORTSAT** Alfabetisk liste over kernebegreber i systematiske oversigtsartikler.

Begreb	Definition/forklaring
Netværksmetaanalyse	En variant af metaanalyse, hvor man sammenligner effekten af $\geq 3$ behandlinger samtidigt med henblik på at kunne få en samlet rangering af effekten af behandlingerne Resultaterne for forskellen mellem 2 behandlinger i netværket vil også bero på indirekte sammenligninger, f.eks. hvis 2 behandlinger hver især sammenlignes med placebo, men ikke med hinanden
Paraplyoversigt eller oversigt af oversigter (eng.: umbrella review)	En variant af systematisk oversigt, hvor man samler og beskriver beslægtede systematiske oversigter
PICO	En engelsk forkortelse, som bruges til at strukturere et terapeutisk forskningsspørgsmål: population, intervention, comparator (kontrolgruppe) og outcome (effekt mål) [3] Et effektmål karakteriseres bl.a. af målemetode, f.eks. visuel analog skala, og tidspunkt for vurdering, f.eks. efter 1 år
PROSPERO	En database til forhåndsregistrering af systematiske oversigter
Publikationsbias	Se »Rapporteringsbias«
Random effects-metaanalyse	En metaanalyse, hvor det antages, at variation mellem studierne skyldes både måleusikkerhed og statistisk heterogenitet Det vil sige, at studierne antages at kunne have lidt forskellige underliggende sande effekter, hvor det vægtede gennemsnit vil være et gennemsnit af disse forskellige sande effekter
Rapporteringsbias	Bruges om risikoen for, at resultaterne fra upublicerede studier (publikationsbias) eller ikke rapporterede resultater (bias pga. selektiv ikke rapportering) systematisk afviger fra de analyserede studieresultater
Risiko for bias	Risikoen for, at resultatet i et studie afviger systematisk fra den sande effekt I randomiserede forsøg kan dette være pga. f.eks. utilstrækkelig randomisering, manglende blinding eller frafald Kan vurderes med standardiserede værktøjer såsom RoB 2 for randomiserede forsøg [4]
RoB 2	Et standardiseret værktøj til at hjælpe til at vurdere af risiko for bias (se »Risiko for bias«) i randomiserede forsøg [4]
ROBIS	Et standardiseret værktøj til at hjælpe med at vurdere risiko for bias (se »Risiko for bias«) i systematiske oversigter [5]
Statistisk heterogenitet	Angiver, hvor forskelligartede de enkelte studieresultater er, ud over hvad der kunne forventes af den statistiske måleusikkerhed Kan skyldes kliniske og metodologiske forskelle imellem studierne Kvantificeres f.eks. ved $I^2$ eller $\tau^2$ , hvor førstnævnte udtrykker procentdelen af variationen, der skyldes forskelle mellem studierne, og dermed er højere ved høj statistisk heterogenitet Statistisk heterogenitet kan forsøges forklaret ved subgruppeanalyser (se »Subgruppeanalyse«) eller metaregression (se »Metaregression«)
Subgruppeanalyse	En analyse, hvor man i metaanalysen opdeler studierne i $\geq 2$ kategorier, f. eks. type af kontrolgruppe, for at undersøge, om der er forskelle mellem disse kategorier, der kan forklare evt. heterogenitet En evt. forskel i effekt mellem kategorierne kaldes også effektmodifikation
Systematisk oversigt (eng.: systematic review)	Et studie, der systematisk identificerer, analyserer og vurderer studier med samme forskningsspørgsmål I oversigten kan også indgå en metaanalyse (se »Metaanalyse«)
Tiltro til evidensen	En samlet vurdering af, hvor sikker man er på, at det fundne resultat er tæt på den sande effekt, ved at tage højde for begrænsninger i evidensen, f.eks. statistisk præcision, heterogenitet, generaliserbarhed, risiko for bias og publikationsbias Kan vurderes med f.eks. GRADE (se »GRADE«)

## CASE

En 51-årig kvinde henvender sig i din praksis pga. skuldersmerter, som er tilkommet over nogle uger. Du undersøger hende og vurderer, at der er tale om subakromielt smertesyndrom (SAPS), hvorfor du henviser hende til fysioterapi. Efter tre mdr. vender hun tilbage, da smerterne ikke er blevet bedre. Hun spørger til en pladsskabende operation i skulderen.

Du finder en systematisk oversigt fra Cochrane (Tabel 1) [6]. I den er følgende forskningsspørgsmål undersøgt: Hos patienter med SAPS i tre mdr. hvad er effekten af subakromiel dekompression, dvs. en pladsskabende

operation, sammenlignet med ikkekirurgisk behandling (f.eks. træning), når man måler på smerteintensitet et år senere? Oversigtens forfattere konkluderer, at der formentlig ikke er nogen relevant forskel i effekt mellem dekompression og træning.

## HVORFOR SYSTEMATISKE OVERSIGTER?

Det er ofte en udfordring at danne sig et overblik over effekten af en given behandling. Måske er det svært at finde relevante studier. Måske har studierne modstridende resultater eller er af tvivlsom kvalitet. En systematisk oversigt er et studie af studier, hvor forskere systematisk identificerer, analyserer og vurderer disse »primærstudier«. Oversigten er dermed et redskab til at skabe det manglende overblik [7].

Vores eksempel på en systematisk oversigt inkluderer randomiserede forsøg til at vurdere behandlingseffekt, og gennemgangen nedenfor er rettet mod denne slags systematiske oversigter [3]. Mange elementer er dog også relevante for systematiske oversigter med andre typer af kliniske spørgsmål, og som dermed inkluderer andre studiedesign, f.eks. diagnostiske akkuratessesstudier [8].

## CENTRALE STUDIEKOMPONENTER

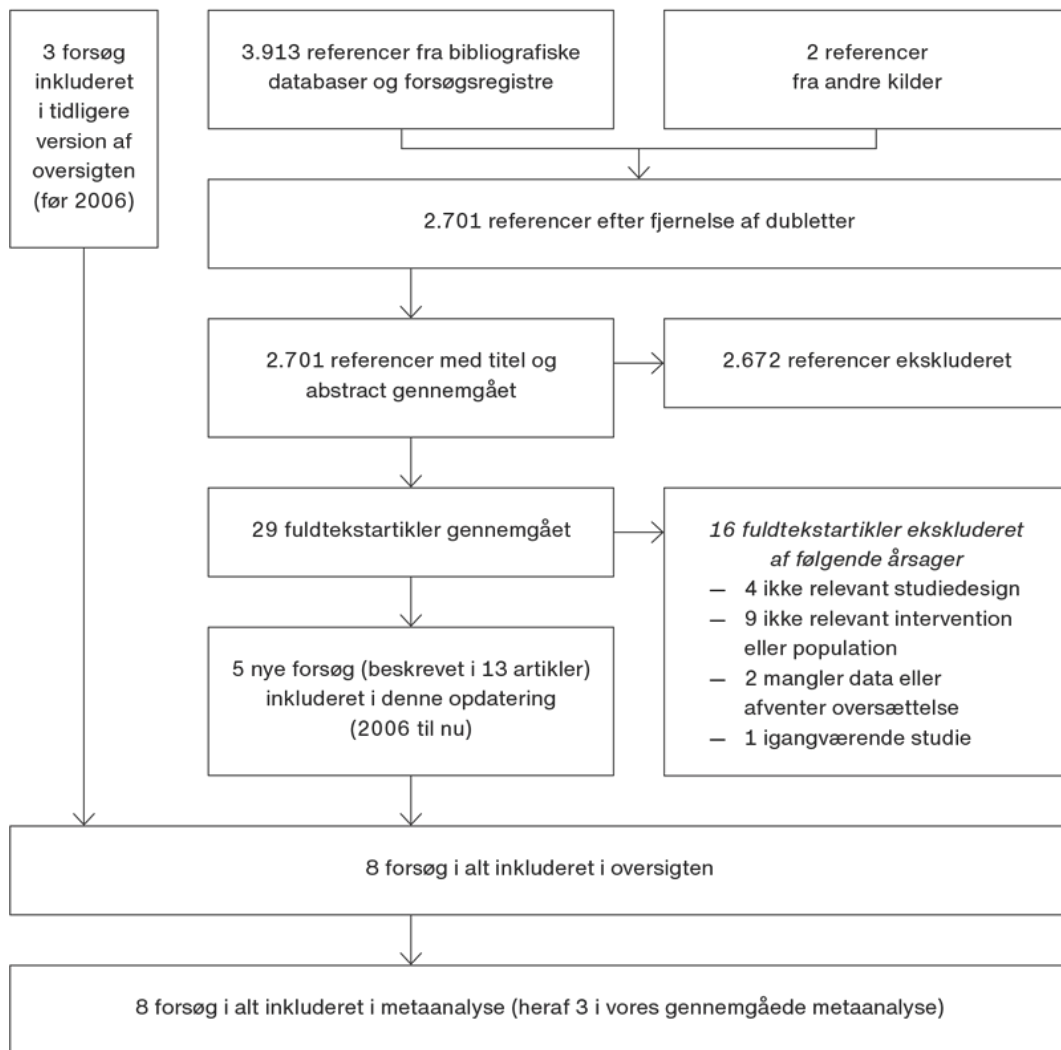
En systematisk oversigt består af en række centrale studiekomponenter (**Tabel 2**). Overordnet vil forskerne ud fra et forskningsspørgsmål og inklusions- og eksklusionskriterier søge efter og udvælge relevante studier (identifikation), indsamle og gennemgå studieresultater f.eks. i en metaanalyse (analyse) og til sidst vurdere tiltroen til den samlede evidens (vurdering). Resultaterne fra de forskellige dele vil typisk blive præsenteret i karakteristiske figurer, f.eks. udvælgelse af studier i et flowdiagram (**Figur 1**) og en eventuel metaanalyse i et forest plot (**Figur 2**).



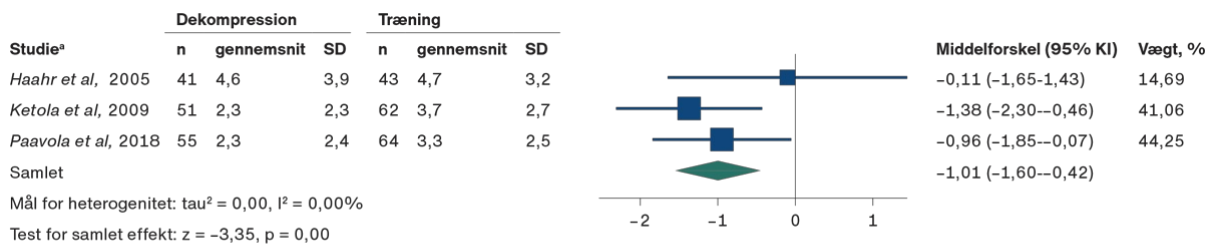
**TABEL 2** De centrale komponenter, som indgår i en systematisk oversigt.

Komponent	Definition/forklaring
Forskningsspørgsmål	Et terapeutisk forskningsspørgsmål vil ofte formuleres ud fra forkortelsen PICO (population, intervention, comparator, outcome) Forskningsspørgsmålet er styrende for de efterfølgende komponenter
Inklusions- og eksklusionskriterier	Udspecificering af særlige krav til studierne såsom population, f.eks. personer over 18 år, eller studiedesign, f.eks. kun randomiserede forsøg
Litteratursøgning og udvælgelse af studier	Formulering af søgestrategi i bibliografiske databaser og i andre kilder samt den efterfølgende proces med at udvælge, hvilke af de fremkomne søgeresultater der opfylder inklusions- og eksklusionskriterier Typisk vil udvælgelsen bestå af to trin: screening af titler og abstracts for oplagte eksklusioner, f.eks. forkert intervention, hvorefter de resterende potentielle kandidater gennemlæses i fuldtekst, og man danner den endelige liste af inkluderede studier
Dataindsamling og kritisk vurdering	Indsamling af data for de inkluderede studier, f.eks. detaljer om population og intervention samt studieresultater Dernæst kritisk vurdering af risiko for bias, typisk vha. et standardiseret værktøj Til randomiserede forsøg kan f.eks. bruges RoB 2, som tager stilling til bl.a. utilstrækkelig randomisering og manglende blinding [4]
Syntese af studieresultater	Opsamling af resultaterne fra de enkelte studier i form af en metaanalyse, dvs. kvantitativ syntese, eller alternativt en kvalitativ opsummering, dvs. kvalitativ syntese Evt. også undersøgelse af årsager til heterogenitet vha. subgruppeanalyser og metaregression
Vurdering af tiltroen til evidensen	En samlet vurdering af, hvor sikker man er på resultaterne GRADE er et hyppigt anvendt værktøj til dette [2]

**FIGUR 1** Flowdiagram for en systematisk oversigt. Diagrammet viser – let forenklet – resultaterne for udvælgelse af forsøg i den gennemgåede systematiske oversigt, som indeholder et mere detaljeret flowdiagram [6]. Referencerne fra alle søgninger samles, og dubletter fjernes (dubletter optræder f.eks., hvis samme reference er fundet i flere databaser). Herefter gennemgås først titler og abstracts, hvorved man kan ekskludere de oplagt irrelevante referencer. Derefter gennemlæses potentielle kandidater i fuldtekst, og man ender derefter med den endelige liste af inkluderede forsøg til oversigten. Alle inkluderede forsøg i vores eksempel havde tilstrækkelige data til at indgå i metaanalyse (heraf tre i vores gennemgåede metaanalyse). Da vores eksempel er en opdatering, tilkommer også de tidligere inkluderede forsøg fra den gamle oversigt (fra kassen øverst til venstre).



**FIGUR 2** Forest plot for metaanalyse af smerteintensitet efter et år for dekompression sammenlignet med træning [6]. I figuren ses resultater både med tal og grafisk for de enkelte studier og for det samlede resultat af metaanalysen. For hvert studie ser man resultatet i smertescorere for intervention og kontrol (for et kontinuert effektmål: antal personer, gennemsnit og standardafvigelse (SD)) samt effekten mellem grupperne (middelforskel) med præcision i form af 95% KI. Desuden vises den udregnede vægt, hvormed hvert studie indgår i metaanalysens vægtede gennemsnit. Dette gennemsnit præsenteres med et 95% KI og grafisk som en diamant til højre. Den samlede middelforskel var  $-1,01$  (95% konfidens-interval (KI):  $-1,60$ – $-0,42$ ) (negativ forskel betød mindre smerteintensitet ved operation end efter træning); denne forskel var statistisk signifikant, men ikke sikkert klinisk relevant (se artiklen for uddybning). Figuren kan også præsentere forskellige mål for statistisk heterogenitet, f.eks.  $I^2$  eller  $\tau^2$ . Derudover vil figuren ofte præsentere den statistiske model, her en random effects-metaanalyse med den såkaldte DerSimonian-Laird-metode.



Random-effects DerSimonian-Laird model

a) Vedr. bibliografiske data: Se [6].

## KRITISK LÆSNING AF EN SYSTEMATISK OVERSIGT

Der udgives mange systematiske oversigter, desværre også flere, der ikke er af tilstrækkelig god kvalitet [9], hvilket skærper behovet for veludførte systematiske oversigter. Cochraneoversigter f.eks. følger strenge metodologiske krav og vil ofte være af høj kvalitet. Som læser er det dog vigtigt at være kritisk og selvstændigt vurdere, om en given oversigt er pålidelig, altså om dens resultater og konklusioner er til at stole på. Vi foreslår derfor, at du stiller følgende ti spørgsmål:

### Er oversigten det mest relevante studie til at besvare forskningsspørgsmålet?

Forskningsspørgsmålet kan være besvaret af andre bedre eller nyere oversigter (eller af nyere forsøg), som kan være relevante for læseren at inddrage og sammenligne med [10].

Vores eksempel [6] er en opdatering af oversigt fra 2008 [11] foranlediget af nye randomiserede forsøg. Foruden denne nyeste Cochraneoversigt kan man også se på en lignende systematisk oversigt, der er udført og udkommet parallelt med denne [12].

### Er oversigtens metoder og analyser fastlagt på forhånd?

Oversigten kan referere til en publiceret protokol eller en offentlig forhåndsregistrering af oversigten, f.eks. i databasen PROSPERO (Tabel 1), og disse kan hjælpe med at sikre gennemsigtighed og overensstemmelse mellem protokol og den færdige oversigt. Metodologiske valg, f.eks. analysemodel, bør være fastlagt på forhånd og ikke truffet på baggrund af primærstudiernes resultater [13]. En oversigt uden en protokol eller registrering bør derfor fortolkes med stor varsomhed [14].

I vores eksempel er den tidligere version forudgået af en publiceret protokol [15], og forfatterne af opdateringen har begrundet protokolafvigelser.

### Har oversigten en dækkende og opdateret søgestrategi?

Der findes vejledninger til vurdering af oversigtens søgning [16]. Bl.a. bør forfatterne have søgt i flere relevante

databaser, f.eks. MEDLINE og Embase, med korrekt opsatte søgestrategier, f.eks. såkaldte bloksøgninger (Tabel 1). Oversigten bør også have ledt i andre informationskilder, f.eks. grå litteratur (Tabel 1) [17], referencelister og artikler, der citerer de inkluderede studier. For ikke at overse nye studier bør søgningen være nyligt opdateret, f.eks. 6-12 mdr. inden udgivelse, gerne endnu tættere på, hvis det drejer sig om et felt i hastig udvikling (f.eks. behandlinger under COVID-19-pandemien).

Vores eksempel har en veludført søgestrategi opdateret få måneder inden publicering, og forfatterne har søgt i flere bibliografiske databaser, inklusive MEDLINE og Embase, i forsøgsregistre ClinicalTrials.gov og World Health Organization ICTRP samt i referencelister i de inkluderede studier.

## **Er udvælgelsen af studier og dataindsamling fra de inkluderede studier udført uafhængigt af mindst to personer?**

I udarbejdelsen af oversigten kan der ske fejl f.eks. pga. overseelser ved udvælgelse af studier, dataindsamling eller vurdering af risiko for bias. Dette kan mindskes ved, at to personer udfører de samme opgaver uafhængigt af hinanden og dernæst afklarer uenigheder, eventuelt ved en tredje persons mellemkomst [18, 19], hvilket også er gjort i vores eksempel.

## **Tager oversigten højde for risiko for bias i de inkluderede studier samt for risikoen for rapporteringsbias (f.eks. publikationsbias)?**

En god systematisk oversigt vil også vurdere og inddrage risiko for bias (Tabel 1), typisk vha. et standardiseret værktøj som Risk of Bias (RoB) 2 til randomiserede forsøg [4]. Hvis mange inkluderede studier har høj risiko for bias, kan effekten af en behandling blive systematisk fejlvurderet [20], og denne risiko for bias vil også videreføres til oversigtens resultater.

Desuden kan resultater fra upublicerede studier eller studier, som har målt, men ikke rapporteret et bestemt effektmål, give anledning til rapporteringsbias (Tabel 1), hvis resultaterne ikke har kunnet skaffes på anden vis, f.eks. ved kontakt til forfatterne [21].

I vores eksempel har forfatterne brugt forløberen til [4] og i fortolkningen af deres resultater taget forbehold for, at studierne har høj risiko for bias pga. manglende blinding. For at mindske risikoen for rapporteringsbias har forfatterne søgt i forsøgsregistre efter upublicerede forsøg og undersøgt protokoller for ikke-rapporterede effektmål.

## **Er de inkluderede studier tilstrækkeligt sammenlignelige til, at en metaanalyse er meningsfuld, og er en eventuel betydende statistisk heterogenitet forklaret?**

I syntesen af de enkelte studieresultater kan forfatterne lave en metaanalyse til at udregne et vægtet gennemsnit (Tabel 1 og Figur 2), eller de kan lave en kvalitativ opsummering. Hvis studierne vurderes tilstrækkeligt klinisk og metodologisk sammenlignelige, vil en metaanalyse give god mening [22]. Hvis metaanalysen viser betydelig statistisk heterogenitet (Tabel 1), vil effekten for forskellige populationer kunne afvige væsentligt fra metaanalysens vægtede gennemsnit, som altså skal tages med forbehold. Forfatterne kan forsøge at forklare heterogeniteten i (planlagte) subgruppeanalyser eller ved metaregression (Tabel 1).

I vores eksempel har forfatterne udført en metaanalyse af tre sammenlignelige forsøg med træning som kontrol, og de har ikke konstateret statistisk heterogenitet, hvilket også kvalitativt kan ses ved det store overlap mellem de tre forsøgs konfidensintervaller (KI) i Figur 2.

## **Vil konklusionen om behandlingens effekt afhænge af, om den nedre og øvre grænse af konfidensintervallet var det sande effektestimat?**

Resultatet af forfatternes metaanalyse er præsenteret i et forest plot i Figur 2. Samlet var smerteintensiteten på en skala med 0-10 point gennemsnitligt 1,01 point lavere ved operation sammenlignet med træning, og dette var



statistisk signifikant forskelligt fra 0 points forskel (95% KI: 1,60 point lavere til 0,42 point lavere,  $p < 0,001$ ). Det væsentlige er dog, om en sådan beskedent forskel er klinisk relevant (Tabel 1). Forfatterne har på forhånd defineret en forskel på 1,5 point som den mindste klinisk relevante forskel på baggrund af en anden systematisk oversigt, der har undersøgt netop dette [23]. Kigger man på hele bredden af 95% KI, så vil en mulig forbedring på 1,60 point (nedre grænse for KI) lige netop være klinisk relevant (da den er større end 1,5 point), mens 1,01 point (punkttestimatet) eller 0,42 point (øvre grænse af KI) ikke vil være det. Denne statistiske usikkerhed – aflæst i KI – er altså en begrænsning og indgår også i forfatterens fortolkning af resultatet.

### **Kan resultaterne generaliseres til den kliniske kontekst?**

I forhold til ekstern validitet kan visse afvigelser fra den kliniske kontekst betyde, at resultaterne måske ikke kan generaliseres, bl.a. hvad angår population, intervention, kontrol og effektmål [24]. I vores eksempel vurderes studiepopulationerne til at være relevante for patienter ligesom vores case, f.eks. når man ser på alder og symptomvarighed.

### **Har oversigtens konklusioner sikkert belæg i resultaterne og er de uden »spin«?**

Forfatterne bør fortolke deres resultater i lyset af eventuelle begrænsninger, som f.eks. risiko for bias. Det vil sige, at de bør vurdere tiltroen til evidensen (Tabel 1). Dette kan f.eks. foregå vha. et værktøj som GRADE [2]. Samtidig bør læseren holde øje med »spin« i konklusionerne, hvor en gavnlig effekt uforbeholdent fremhæves, selvom resultatet havde et bredt KI. I vores eksempel har forfatterne brugt GRADE til at konkludere på afbalanceret vis og uden spin.

### **Er oversigten kommercielt finansieret, eller har forskerne andre interessekonflikter, der kan have påvirket udarbejdelsen af oversigten?**

Kommerciel involvering i oversigter er associeret med mere positive konklusioner, mens indflydelsen på selve resultaterne er mere uklare [25]. Cochraneoversigter som vores eksempel er ikke kommercielt støttede, og forfatterne rapporterede ikke nogen relevante interessekonflikter.

## **DESIGNVARIANTER**

I det ovenstående har vi taget udgangspunkt i en klassisk systematisk oversigt med metaanalyse. Eksempler på andre typer af oversigter og metaanalyser findes i Tabel 1, nemlig kortlægningsoversigter (på engelsk scoping reviews), hurtigoversigter (rapid reviews), paraplyoversigter (umbrella reviews), netværksmetaanalyser og metaanalyser med individuelle patientdata.

## **RAPPORTERINGSRETNINGSLINJER OG VÆRKTØJER TIL KRITISK VURDERING**

For at kunne vurdere et studie på et ordentligt grundlag er det nødvendigt, at forskningsartiklen indeholder visse essentielle oplysninger. For systematiske oversigter findes rapporteringsretningslinjen PRISMA 2020 [26], som kan klargøre for både forfatter og læser, hvad der som minimum bør rapporteres.

De ti spørgsmål ovenfor giver en simpel tilgang til kritisk vurdering af en systematisk oversigt. Til den grundigere gennemgang findes standardiserede værktøjer som AMSTAR 2 og ROBIS [1, 5], der også uddyber mange af spørgsmålene.

## **FRA EVIDENS TIL HANDLING**

Vi har ovenfor fokuseret på ét forskningsspørgsmål med én slags kontrol (ikkekirurgisk behandling, f.eks.

træning) og ét effektmål (smerte efter et år). Oversigten inddrager imidlertid også forsøg, hvor kontrolgruppen fik placebokirurgi, og den kigger på andre effektmål, f.eks. smerte efter seks mdr., funktion og skadevirkninger. Dette kan bidrage til det samlede billede af balancen mellem gavnlige og skadelige effekter, og en sådan afvejning indgår ved udarbejdelsen af evidensbaserede anbefalinger, f.eks. kliniske retningslinjer [27]. For vores patientcase findes der to kliniske retningslinjer [28, 29], som også inddrager den gennemgåede oversigt, og ud fra en samlet vurdering af evidensen fraråder de begge operation.

**Korrespondance** *David Ruben Teindl Laursen*. E-mail: [dlaursen@health.sdu.dk](mailto:dlaursen@health.sdu.dk)

**Antaget** 20. februar 2024

**Publiceret på ugeskriftet.dk** 25. marts 2024

**Interessekonflikter** ingen. Forfatterens ICMJE-formularer er tilgængelige sammen med artiklen på [ugeskriftet.dk](https://ugeskriftet.dk)

**Taksigelse** *Stefan Kern Nielsen, Emeka Onwuneme og Kasthuren Ranjan* takkes for kommentarer til manuskriptet

**Referencer** findes i artiklen publiceret på [ugeskriftet.dk](https://ugeskriftet.dk)

**Artikelreference** *Ugeskr Læger* 2024;186:V09230616

**doi** [10.61409/V09230616](https://doi.org/10.61409/V09230616)

**Open Access** under Creative Commons License [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)

## SUMMARY

### Critical reading of systematic review articles

A systematic review provides an overview of primary studies investigating a given research question, e.g., the effect of a certain treatment. Individual study results are sometimes synthesised in a meta-analysis. A critical reader should consider whether the systematic review is relevant and reliable, e.g., does it follow a protocol, address the risk of bias, and consider potential heterogeneity. PRISMA 2020 guideline recommends a minimum set of items that should be reported in a systematic review article, and AMSTAR 2 and ROBIS are tools for critical appraisal of systematic reviews.

## REFERENCER

1. Shea BJ, Reeves BC, Wells G et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358:j4008. doi: <https://doi.org/10.1136/bmj.j4008>
2. Guyatt G, Oxman AD, Akl EA et al. GRADE guidelines: 1. Introduction - GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64(4):383-94. doi: <https://doi.org/10.1016/j.jclinepi.2010.04.026>
3. Higgins JPT, Thomas J, Chandler J et al, red. *Cochrane handbook for systematic reviews of interventions*. Version 6.4 (updated August 2023). Cochrane, 2023. [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook) (5. jul 2023).
4. Sterne JAC, Savovi&; J, Page MJ et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898. doi: <https://doi.org/10.1136/bmj.l4898>
5. Whiting P, Savovi&; J, Higgins JPT et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. 2016;69:225-234. Doi: <https://doi.org/10.1016/j.jclinepi.2015.06.005>
6. Karjalainen TV, Jain NB, Page CM et al. Subacromial decompression surgery for rotator cuff disease. *Cochrane Database Syst Rev*. 2019;1(1):CD005619. doi: <https://doi.org/10.1002/14651858.CD005619.pub3>
7. Hróbjartsson A. Studier af andre studier: systematiske oversigter og metaanalyser. I: Hróbjartsson A, Lundh A, red. *Evidensbaseret medicin og klinisk forskningsmetode*. 1. udg. Munksgaard, 2022:263-297.
8. Deeks J, Bossuyt P, Leeflang M, Takwoingi Y, red. *Cochrane handbook for systematic reviews of diagnostic test accuracy*. Version 2.0 (updated July 2023). Cochrane, 2023. doi: <https://doi.org/10.1002/14651858.ED000163>

9. Ioannidis JPA. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q*. 2016;94(3):485-514. doi: <https://doi.org/10.1111/1468-0009.12210>
10. BMJ Best Practice. Multiple systematic reviews on the same question: some considerations, 2022. <https://bestpractice.bmj.com/info/toolkit/learn-ebm/multiple-systematic-reviews-on-the-same-question/> (24. nov 2023).
11. Coghlan JA, Buchbinder R, Green S et al. Surgery for rotator cuff disease. *Cochrane Database Syst Rev*. 2008;2008(1):CD005619. doi: <https://doi.org/10.1002/14651858.CD005619.pub2>
12. Lähdeoja T, Karjalainen T, Jokihaara J et al. Subacromial decompression surgery for adults with shoulder pain: a systematic review with meta-analysis. *Br J Sports Med*. 2020;54(11):665-673. doi: <https://doi.org/10.1136/bjsports-2018-100486>
13. Tricco AC, Cogo E, Page MJ et al. A third of systematic reviews changed or did not specify the primary outcome: a PROSPERO register study. *J Clin Epidemiol*. 2016;79:46-54. doi: <https://doi.org/10.1016/j.jclinepi.2016.03.025>
14. Sandau N, Aagaard TV, Hróbjartsson A et al. A meta-epidemiological study found that meta-analyses of the same trials may obtain drastically conflicting results. *J Clin Epidemiol*. 2023;156:95-104. doi: <https://doi.org/10.1016/j.jclinepi.2023.02.002>
15. Coghlan JA, Buchbinder R, Green S, Bell SN. Surgery for rotator cuff disease. *Cochrane Database Syst Rev*. 2006;2006(1):CD005619. doi: <https://doi.org/10.1002/14651858.CD005619>
16. Frandsen TF, Dyrvig AK, Christensen JB et al. En guide til valide og reproducerbare systematiske litteratursøgninger. *Ugeskr Læger*. 2014;176:V02130141.
17. Hopewell S, McDonald S, Clarke M, Egger M. Grey literature in meta-analyses of randomized trials of health care interventions. *Cochrane Database Syst Rev*. 2007;2007(2):MR000010. doi: <https://doi.org/10.1002/14651858.MR000010.pub3>
18. Waffenschmidt S, Knelangen M, Sieben W et al. Single screening versus conventional double screening for study selection in systematic reviews: a methodological systematic review. *BMC Med Res Methodol*. 2019;19(1):132. doi: <https://doi.org/10.1186/s12874-019-0782-0>
19. Robson RC, Pham B, Hwee J et al. Few studies exist examining methods for selecting studies, abstracting data, and appraising quality in a systematic review. *J Clin Epidemiol*. 2019;106:121-135. doi: <https://doi.org/10.1016/j.jclinepi.2018.10.003>
20. Page MJ, Higgins JPT, Clayton G et al. Empirical evidence of study design biases in randomized trials: systematic review of meta-epidemiological studies. *PloS One*. 2016;11(7):e0159267. doi: <https://doi.org/10.1371/journal.pone.0159267>
21. Page MJ, Sterne JAC, Boutron I et al. ROB-ME: a tool for assessing risk of bias due to missing evidence in systematic reviews with meta-analysis. *BMJ*. 2023;383:e076754. doi: <https://doi.org/10.1136/bmj-2023-076754>
22. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to meta-analysis. 2nd ed. Wiley, 2021. doi: <https://doi.org/10.1002/9781119558378>
23. Hao Q, Devji T, Zeraatkar D et al. Minimal important differences for improvement in shoulder condition patient-reported outcomes: a systematic review to inform a BMJ Rapid Recommendation. *BMJ Open*. 2019;9(2):e028777. doi: <https://doi.org/10.1136/bmjopen-2018-028777>
24. Guyatt GH, Oxman AD, Kunz R et al. GRADE guidelines: 8. Rating the quality of evidence-indirectness. *J Clin Epidemiol*. 2011;64(12):1303-10. doi: <https://doi.org/10.1016/j.jclinepi.2011.04.014>
25. Hansen C, Lundh A, Rasmussen K, Hróbjartsson A. Financial conflicts of interest in systematic reviews: associations with results, conclusions, and methodological quality. *Cochrane Database Syst Rev*. 2019;8(8):MR000047. doi: <https://doi.org/10.1002/14651858.MR000047.pub2>
26. Page MJ, McKenzie JE, Bossuyt PM et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. doi: <https://doi.org/10.1136/bmj.n71>
27. Schroll JB, Jeppesen BT, Møller MH et al. Kritisk læsning af kliniske retningslinjer. *Ugeskr Læger*. 2024;186(6):V09230579. doi: <https://doi.org/10.61409/V09230579>
28. Vandvik PO, Lähdeoja T, Arden C et al. Subacromial decompression surgery for adults with shoulder pain: a clinical practice guideline. *BMJ*. 2019;364:l294. doi: <https://doi.org/10.1136/bmj.l294>
29. Sundhedsstyrelsen. National klinisk retningslinje for behandling af patienter med subakromielt smertesyndrom i skulderen (impingement syndrom/rotator-cuff syndrom), 2021. [https://www.sst.dk/-/media/Udgivelser/2021/NKR\\_skulder/NKR-05-skuldresmerter-retningslinje-final.ashx](https://www.sst.dk/-/media/Udgivelser/2021/NKR_skulder/NKR-05-skuldresmerter-retningslinje-final.ashx) (29. aug 2023).