

Bilaga 6 Stegen i den systematiska översikten

Urvalskriterier

Urvalskriterier formuleras och beskrivs för att på ett transparent sätt avgöra och avgränsa vilken forskning som är relevant för översiktens syften och frågeställningar. Urvalskriterierna består av inklusions- och exklusionskriterier som anger villkor för att data i en studie ska accepteras eller inte i den systematiska översikten. Frågeställningar som avser effekt struktureras enligt PICO-modellen, och villkor för inkludering och exkludering beskrivs i förhållande till de fyra komponenterna population eller personer (P), insats (I), jämförelseinsats (eng. control (C)) och utfall (eng. outcome(O)). Frågor om upplevelser och erfarenheter av insatser struktureras enligt SPICE-modellens fem komponenter Sammanhang (S), Perspektiv (P), Insats (I), Jämförelseinsats (eng. control (C)) och Utvärdering (eng. Evaluation (E)).

Litteratursökning

Med alla, eller delar av, dessa kriterier som utgångspunkt formuleras en strategi för litteratursökning med syfte att identifiera samtliga vetenskapliga studier som kan vara aktuella för översikten.

Relevansgranskning

Identifierade studier bedöms därefter i relation till urvalskriterierna samt syften och frågeställningar för att avgöra vilka studier som är faktiskt relevanta.

Relevansgranskningen genomförs av två granskare oberoende av varandra och sker i två steg. Studierna relevans bedöms först med stöd av titel, sammanfattning och nyckelord. Alla studier som av minst en granskare bedöms som möjligt relevanta läses därefter i ett andra steg igenom i fulltext för att bedöma om de

faktiskt är relevanta. För att en studie ska bedömas som faktiskt relevant krävs ett gemensamt beslut av båda granskarna.

Risk för systematiska fel

De studier som slutligen bedöms som faktisk relevanta i förhållande till urvalskriterier, syften och frågeställningar granskas med avseende på risk för systematiska fel. För kvantitativa primärstudier uttrycks risken för systematiska fel med termen *risk för snedvridning av resultat* (eng. Risk of Bias (RoB)), medan motsvarande term för primärstudier baserade på kvalitativa data är *metodologiska brister*. Även detta görs av två granskare oberoende av varandra, och bedömningen görs med stöd av SBU:s granskningsmallar framtagna för respektive studietyp, se SBU:s metabok [1].

Analys och syntes

Resultaten från studier som har en acceptabel risk för systematiska fel vägs samman i analys och syntes om de är tillräckligt lika, annars beskrivs de deskriptivt. För att översiktens läsare ska få en överblick av den data som ligger till grund för analys och syntes extraheras och sammanställs, som en del av resultatet, sådan data i tabeller. Tabellerna är skrivna på engelska för att innehållet ska vara tillgängligt för intresserade från andra länder.

Data som extraheras kan vara:

- författare
- land där studien är gjord
- studiens syfte/frågeställning
- studiens design
- teori eller ansats, samt forskarnas antal/kompetens
- sammanhang där studien är genomförd (setting)
- studieperiod/uppföljningstid
- beskrivning av deltagare som ingick i studien
- storlek på eventuella bortfall i studien
- beskrivning av utfallsmått och metoder för att mäta utfall

- beskrivning av den eller de insatser och jämförande insatser som utvärderats i studien
- datainsamlingsmetod
- dataanalysmetod

Tillförlitlighet i det samlade vetenskapliga underlaget

I nästa steg bedöms och klassificeras hur tillförlitligt det samlade vetenskapliga underlaget är. Som stöd för bedömningen av resultat från kvantitativa studier använder SBU det internationellt utarbetade systemet GRADE (The Grading of Recommendations Assessment, Development and Evaluation) [2]. Tillförlitligheten för fynd från kvalitativa studier bedöms med stöd av GRADE-CERQual [3].

GRADE syftar till att på ett strukturerat och transparent sätt bedöma osäkerheter och risker i resultatet baserade på kvantitativa data. Resultatet granskas ur olika synvinklar och utgångspunkterna för bedömning av resultatens tillförlitlighet är fyra riskområden (Faktaruta 1). GRADE tillämpas på samma sätt oavsett om resultaten är sammanvägda eller narrativa. Samma tillämpning av GRADE gäller också oavsett om underlaget är omfattande eller litet, det vill säga består av en enda studie eller ett fåtal små studier. Brister i samstämmighet är dock endast relevant att bedöma om det finns mer än en studie. GRADE-bedömning görs per utfallsmått.

Faktaruta 1 Bedömning av resultatens tillförlitligt utifrån fem riskområden.**Riskområde 1 - Risk för snedvridning**

Detta riskområde gäller hur stor risken är att det (sammanvägda) estimatet påverkas av brister i studierna.

Riskområde 2 – Bristande samstämmighet

Med bristande samstämmighet, heterogenitet, avses att studierna visar olika resultat.

Riskområde 3 – Bristande precision

När vi tittar på precision bedömer vi bredden och läget konfidensintervallet för det sammanvägda resultatet. Bristande precision kan också handla om ett lågt antal händelser eller ett lågt antal deltagare.

Riskområde 4 – Bristande överförbarhet

Överförbarhet innebär att resultatet från studierna kommer att vara likartat för det sammanhang som forskningsfrågan avser. Brister i överförbarheten kan relateras till skillnader i population, skillnader i intervention, skillnader i utfallsmått samt indirekta jämförelser.

Riskområde 5 – Publikationsbias

Publikationsbias är sannolikt mycket vanligt förekommande, men det är svårt att bedöma hur allvarlig risken är. Det finns flera metoder som kan ge en fingervisning om att det saknas studier men det behövs indicier från mer än en metod för att göra avdrag. Till skillnad från övriga riskområden i GRADE kan man inte dra av med mer än ett steg. Riktlinjerna för GRADE rekommenderar att man överväger att göra avdrag med ett steg om underlaget enbart består av små studier. Om de dessutom är sponsrade av företag eller om prövarna har någon annan form av intressekonflikter ökar risken för publikationsbias.

Med GRADE klassificeras tillförlitligheten som hög, måttlig, låg eller mycket låg. När det helt saknas studier som uppfyller inklusionskriterierna anges ”studier saknas”, utan gradering av tillförlitligheten. Innebörden av de olika domäner som bedöms och de olika nivåerna av tillförlitlighet framgår av Faktaruta 2.

Faktaruta 2 Resultatens tillförlitlighet klassificeras i en av fyra nivåer enligt GRADE.

- Det sammanvägda resultatet har hög tillförlitlighet (⊕⊕⊕⊕)
- Det sammanvägda resultatet har måttlig tillförlitlighet (⊕⊕⊕○)
- Det sammanvägda resultatet har låg tillförlitlighet (⊕⊕○○)
- Det sammanvägda resultatet har mycket låg tillförlitlighet (⊕○○○)
(Det går inte att bedöma om resultatet stämmer)

När det helt saknas studier som uppfyller inklusionskriterierna anges ”studier saknas”, utan gradering av tillförlitligheten.

Det är viktigt att poängtera att graden av tillförlitlighet i det vetenskapliga underlaget inte har bäring på insatsens effektstorlek eller allvarlighetsgrad av risker, utan det är en metod för att uttrycka hur stor säkerhet som finns i underlaget för ett visst resultat i sig.

Resultat för olika utfallsmått och deras grad av tillförlitlighet redovisas i en sammanfattande resultattabell, en så kallad Summary of findings-tabell (SoF-tabell).

Tillförlitligheten för resultat i form av fynd från kvalitativa synteser bedöms med stöd av GRADE-CERQual. Syftet är att på ett transparent sätt bedöma och beskriva hur stor tilltro som beslutsfattare och andra kan fästa på fynden. CERQual definierar tillförlitligheten som en bedömning av i vilken utsträckning fyndet är en rimlig representation av fenomenet. En alternativ formulering är i vilken utsträckning fyndet är ”substansiellt” skiljt från fenomenet. Med det menas att skillnaden är så stor att det påverkar beslutsfattandet.

Med fynd avses resultat från ett analytiskt arbete som, baserat på data från primärstudier, beskriver ett fenomen eller en aspekt av ett fenomen. CERQual är inspirerat av GRADE och har utvecklats i samarbete med GRADE Working Group. CERQual är avsett att fungera som ett strukturerat stöd för bedömningar och tolkningar som kommer att vara subjektiva. CERQual består av fyra riskområden (Faktaruta 3).

Faktaruta 3 Bedömning av tillförlitligheten för fynd från kvalitativa synteser.

Riskområde 1 - Metodologiska begränsningar

Med metodologiska begränsningar avses i vilken utsträckning design och genomförande av studierna påverkar tillförlitligheten.

Riskområde 2 – Relevans

Med relevans avses i vilken utsträckning data från de underliggande studierna är tillämpliga för forskningsfrågan.

Riskområde 3 – Koherens

Med koherens avses att fyndet är väl underbyggt av data från studierna och ger en övertygande förklaring för de mönster som identifieras.

Riskområde 4 - Tillräckliga data

Riskområdet handlar dels om hur rikt data är, dels kvantiteten data.

På samma sätt som med GRADE utgår man från att fyndet är tillförlitligt och gör avdrag för brister som kan påverka tillförlitligheten. Tillförlitligheten klassificeras i fyra nivåer. Faktaruta 4 beskriver hur de olika nivåerna kan tolkas.

Faktaruta 4 Klassificering av tillförlitlighet enligt CERQual.

Nivå	Förklaring
Hög tillförlitlighet ⊕⊕⊕⊕	Det är mycket sannolikt att fyndet är en rimlig representation av fenomenet i fråga
Måttlig tillförlitlighet ⊕⊕⊕○	Det är sannolikt att fyndet är en rimlig representation av fenomenet i fråga
Låg tillförlitlighet ⊕⊕○○	Det är möjligt att fyndet är en rimlig representation av fenomenet i fråga
Mycket låg tillförlitlighet ⊕○○○	Det går inte att avgöra om fyndet är en rimlig representation av fenomenet i fråga

I likhet med GRADE sammanställs fynd, antal underliggande studier och deltagare samt en sammanvägd bedömning av tilltron till fyndet i en SoF-tabell. Motiven till avdrag ska framgå i anslutning till tabellen, till exempel i form av fotnoter.

För ytterligare information om metodiken hänvisar vi till SBU:s metodbok [1].

Referenser

1. SBU. Utvärdering av metoder i hälso- och sjukvården och insatser i socialtjänsten: en metodbok. Stockholm: Statens beredning för medicinsk och social utvärdering (SBU); 2020. [accessed Sep 20 2021]. Available from: <https://www.sbu.se/metodbok>.
2. GRADE. The GRADE working group. [accessed April 27 2021]. Available from: <https://www.gradeworkinggroup.org/>.
3. GRADE-CERQual. The GRADE-CERQual Project Group. [accessed Sep 20 2021]. Available from: <https://www.cerqual.org/>.