

Bilaga 3. Granskningsmallar

Granskningsmallarna är baserade på såväl internationella erfarenheter som synpunkter och erfarenheter från projektledare, expertgrupper och externa rådgivare. Två personer i SBU:s vetenskapliga råd har varit speciellt behjälpliga i arbetet med att ta fram mallarna. Olof Nyrén vid Karolinska Institutet har utvecklat en grundläggande version för granskningsmallar. Den har använts i flera SBU-projekt. Mats Eliasson har utarbetat en version som används i Socialstyrelsens arbete med Nationella Riktlinjer.

Varje studietyp har sina metodproblem. Därför finns det olika granskningsmallar för olika studietyper:

- Randomiserade kontrollerade prövningar
- Kohortstudier med kontrollgrupper
- Fall–kontrollstudier
- Diagnostiska och prognostiska studier
- Systematiska översikter.

För att bedöma den samlade evidensstyrkan enligt GRADE finns också ett arbetsblad (Bilaga 5).

I granskningsmallarna anges poäng för svaren till frågorna. Ju högre poäng desto högre kvalitet. Poängen är specifik för respektive granskningsmall/studietyp. Poängsystemet kan fungera som en vägledning för att göra en samlad bedömning av kvaliteten i varje enskild studie. Projektgrupperna får själva ta ställning till om man vill använda poängsystemet eller inte.

Granskningsmall för randomiserad kontrollerad prövning

1. Extern validitet

1.1 Rekrytering av prövningsdeltagare

Personer som uppfyller inklusionskriterierna och inte utesluts av något exklusionskriterium benämns här "valbara".

- a) Är inklusionskriterierna tydligt formulerade och acceptabla?
 - Ja (= 2)
 - Nej (= 0)

- b) Är exklusionskriterierna tydligt formulerade och acceptabla?
 - Ja (= 2)
 - Nej (= 0)

- c) Framgår det hur många personer som exkluderades och varför?
 - Ja (= 2)
 - Nej (= 0)

- d) Togs valbara patienter in konsekutivt i studien?
 - Ja (= 1)
 - Nej (= 0)

- e) Har antal och orsaker angivits för de valbara patienter som inte kom att delta?
 - Ja (= 2)
 - Nej (= 0)

Kommentar: Det finns även andra frågor som kan hänföras till begreppet extern validitet såsom frågorna under 3.1 och 3.2.

2. Intern validitet

2.1 Tilldelning av åtgärd/intervention/behandling

- a) Är randomiseringsproceduren beskriven?
 - Ja, i detalj (= 3) Gå vidare till fråga 2.1 b och c
 - Ja, delvis (= 1) Gå vidare till fråga 2.1 d
 - Nej (= 0) Gå vidare till fråga 2.1 d

- b) Var det möjligt att manipulera randomiseringen?
 - Nej (t ex ogenomskinliga kuvert, datorgenererad sekvens som finns hos någon annan än prövningspersonalen) (= 0)
 - Ja (t ex om metoden innebär att singla slant eller kasta tärning) (= -1*)

* Minska i detta fall poängsumman med 1 poäng.

- c) Medförde randomiseringen en oförutsägbar helt slumpmässig allokering (dvs fördelning av deltagarna) mellan prövningsgrupperna?
- Ja (= 0)
 - Nej (= -1*)
- d) Exkluderades några deltagare efter randomisering men före åtgärd/intervention/ behandling?
- Nej, inte alls (= 3)
 - Ja, men med liten risk att påverka resultaten (se kommenter) (= 2)
 - Ja, med stor risk att påverka resultaten (se kommentar) (=0)

Kommentar: Är det möjligt att man genom exklusion av försöksdeltagare har påverkat resultaten? Antalet deltagare som här lämnat studien bör sättas i relation till studiens storlek. Om exklusioner är jämnt fördelade över behandlingsarmarna och av skäl som är angivna och godtagbara så bedöms risken att resultaten kan påverkats som liten. Om mer än 5 procent av de randomiserade patienterna exkluderades eller att skälen inte är angivna eller godtagbara så bedöms risken som stor.

2.2 Gruppernas jämförbarhet

- a) Finns en redogörelse för gruppernas fördelning avseende kända faktorer som kan inverka på resultatet (ex "patient characteristics" eller "demographic characteristics")?
- Ja (= 1)
 - Nej (= 0)
- b) Fanns det några väsentliga skillnader mellan grupperna (obalanser) i baslinjedata?
- Nej (= 1)
 - Ja (= 0)
- c) Om det fanns obalanser, gjordes några försök att korrigera dem i den statistiska analysen?
- Behövdes inte (inga obalanser av vikt fanns) (= 1)
 - Ja (= 1)
 - Nej, trots att det behövdes (= 0)

2.3 Blindning (maskering)

- a) Gjordes försök till blindning, dvs att dölja (maskera) i vilken grupp försöksdeltagarna (patienterna) hamnade?
- Nej (= öppen prövning) (= 0)
 - Enkelblindad – deltagarna var blindade (= 2)
 - Enkelblindad – observatör och/eller behandlare var blindade (= 2)
 - Dubbelblind – både deltagarna och behandlare och/eller observatör var blindade (= 3)
 - Dubbelblind + att det framgår av beskrivningen att all resultatbearbetning utfördes innan prövningskoden bröts (= 4)

* Minska i detta fall poängsumman med en poäng.

- b) Fanns någon anledning att tro att blindningen inte fungerade (exempelvis genom viss biverkan av aktiv behandling eller bristande likhet mellan tabletter/kapslar avseende utseende eller smak)?
- Nej (= 1)
 - Ja (= 0)

Kommentar: Fråga b belyser svårigheterna att blinda såväl deltagare som prövare. Det finns flera klassiska exempel på studier där man pga karakteristiska effekter eller biverkningar av interventionen inte lyckades med blindningen. I vissa studier prövar man om blindningen har fungerat genom att låta studiedeltagarna gissa om de fått aktiv behandling eller kontroll.

2.4 Fölsamhet (compliance)

- a) Framgår det i vilken utsträckning deltagarna fullföljde behandlingen?
- Ja (= 1)
 - Nej (= 0) → Gå vidare till fråga 2.5
- b) Var andelen som fullföljde behandlingen acceptabelt* stor?
- Ja (= 2)
 - Nej (= 0)

Kommentar: Kontroll av fölsamheten är särskilt viktig då man inte ser en signifikant effektskillnad mellan interventionsgrupp och kontrollgrupp. En bristande fölsamhet minskar oftast såväl interventionens effekter som bieffekter. Vid en tydligt signifikant effekt av en intervention är kontroll av fölsamheten ofta av mindre betydelse för att lösa problemen med fölsamhet inte var koncentrerad till referensbehandlingsarmen (är tänkbar i placebokontrollerad studie om blindningen varit otillräcklig, alternativt i en "head-to-head"-jämförelse med en referensbehandling om den senare var associerad med mycket biverkningar).

2.5 Bortfall (antalet deltagare som inte har följts upp enligt studieprotokollet)

- a) Finns en redogörelse för hur stort bortfallet är och specificeras orsakerna till bortfallet?
- Ja (= 3)
 - Nej (= 0) → Gå vidare till fråga 2.6
- b) Hur stor är risken att bortfallet kan ha snedvridit resultaten?
- Liten (exempelvis bortfall <10%) (= 3)
 - Medel (exempelvis bortfall 10–19%) (= 2)
 - Stor (exempelvis bortfall 20–29%) (= 0)
 - Mycket stor (exempelvis bortfall ≥30%) → undersökningen bedöms ofta sakna informationsvärde vilket kan innebära exklusion från evidensbedömning

Kommentar: Generellt ökar stora bortfall risken för att resultaten kan vara påverkade av systematiska fel. Det kan dock finnas tillfällen då man bedömer att även ett större bortfall sannolikt är slumpmässigt. Procentsiffrorna får därför ses som riktvärden. Bortfallet varierar

* Acceptabelt kan t ex vara att >80% av deltagarna fick >80% av den föreskrivna totaldosen.

mellan olika tidpunkter i en studie och kan dessutom variera mellan olika effektmått. Bortfallet är ofta större ju längre tid som har gått. Därmed kan kanske data för de sista besöken vara av tveksam validitet, medan resultaten för de första besöken kan vara giltiga.

2.6 Bedömning av biverkningar

- a) Vidtogs acceptabla åtgärder för att samla in och registrera information om biverkningar?
- Ja, systematisk insamling och rapportering (= 4)
 - Ja, begränsad beskrivning av inträffade biverkningar (= 2)
 - Nej (= 0)

2.7 Analys av resultaten

- a) Var det huvudsakliga (primära) effektmåttet (resultatvariabel och mätmetod) definierat i artikeln?
- Ja (= 1)
 - Nej (= 0)
- b) Var slutsatsen i studien baserad på det primära effektmåttet (enligt a ovan)?
- Ja (= 1)
 - Nej (= 0)
- c) Redovisade studien någon förhandshypotes? (Varför skulle just denna behandling vara effektiv?)
- Ja (= 1)
 - Nej (= 0)
- d) Var sekundära effektmått definierade i artikeln?
- Ja (= 1)
 - Nej (= 0)

Kommentar: Frågorna 2.7 a–d belyser vikten av att det studerade effektmåttet har definierats i studieprotokollet som ett primärt eller sekundärt effektmått. Observera att det kan förekomma att studier redovisar resultat för effektmått som inte har definierats i protokollet. För nyare studier finns studieprotokoll eller synopsis tillgängligt på publika webbsidor (exempelvis www.clinicaltrials.gov eller www.controlled-trials.com). Det är bra att som rutinåtgärd göra en kontroll av angiven primär effektvariabel på dessa webbsidor. Man bör iaktta en viss försiktighet vid användande av kombinerade effektmått.

- e) Har det gjorts en ITT ("intention-to-treat")-analys?
- Ja (= 4)
 - Nej (= 0)

Kommentar: ITT-analys innebär att effekten beräknas på data från samtliga randomiserade deltagare i den behandlingsgrupp som de från början allokerades till. Om resultaten är beräknade på annat sätt än med ITT finns det risk för att behandlingseffekten blir över- eller underskattad.

3. Bedömning av resultat och precision

3.1 Bedömning av resultat

- a) Finns det en acceptabel definition av effektmåttet?
- Ja (= 3)
 - Nej (= 0)
- b) Är effektmåttet kliniskt relevant?
- Ja (= 2)
 - Relevansen diskutabel (= 0)
 - Relevans obefintlig → undersökningen bedöms sakna informationsvärde

3.2 Minsta kliniskt relevanta effekt

- a) Var den minsta kliniskt relevanta effekten förhandsdefinierad?
- Ja (= 1)
 - Nej (= 0)
- b) Är den valda minsta kliniska relevanta effekten av rimlig storlek?
- Ja (= 1)
 - Nej (eller ej angiven) (= 0)

3.3 Undersökningens statistiska styrka

- a) Är de överväganden och beräkningar som ligger till grund för urvalsstorleken ("sample size") tydligt beskrivna?
- Ja (= 2)
 - Nej (= 0)
- b) Hur stor blev styrkan, dvs sannolikheten för att en kliniskt relevant effekt skulle kunna påvisas med acceptabel signifikansnivå?
- Styrkan är inte angiven, men det fanns en starkt statistiskt signifikant effekt (= 3)
 - >90% (= 3)
 - 80–89% (= 2)
 - 70–79% (= 1)
 - <70% (= 0)
 - Inte angiven trots att resultatet inte var signifikant (= 0)

Granskningsmall för kohortstudier med kontrollgrupper

Värdering av interventionsmetoders effekt och säkerhet

Värdering av riskfaktorer/riskmarkörers betydelse för att förutsäga sjukdom

1. Övergripande frågor

1.1 Är frågeställningen/hypotesen klart formulerad?

- Ja (= 2)
- Nej (= 0)

1.2 Är utfallen väl definierade?

- Ja (= 3)
- Nej (= 0)

1.3 Är interventionen/metoden/riskfaktorn väl definierad?

- Ja (= 2)
- Ja, delvis (= 1)
- Nej (= 0)

1.4 Rekryterades kohorten på ett acceptabelt sätt?

- Ja (= 2)
- Kan ej värderas (= 0)
- Nej (= 0)

Kommentar: Är kohorten representativ för den definierade populationen?

1.5 Är inklusionskriterier tydligt angivna?

- Ja (= 1)
- Nej (= 0)

1.6 Är exklusionskriterier tydligt angivna?

- Ja (= 1)
- Nej (= 0)

1.7 Var studiedeltagarna konsekutivt identifierade?

- Ja (= 1)
- Nej (= 0)
- Ej angivet (= 0)

Kommentar: Med konsekutivt urval menas att alla patienter som uppfyller inklusionskriterier väljs under en viss fördefinierad tidsperiod. Metodologiskt kan urvalet påverka vilka typer av patienter man får, men framför allt är det viktigt att försöks- och kontrollgrupperna valts ut på samma sätt.

1.8 Valbara deltagare som avböjde att delta: har antal och orsaker angivits?

- Ja (= 1)
- Nej (= 0)

1.9 Mättes interventionen/exponeringen adekvat?

- Ja (= 3)
- Kan ej värderas (= 1)
- Nej (= 0)

Kommentar: Finns risk för mät- eller klassifikationsbias? Användes objektiva eller subjektiva metoder? Var metoderna validerade? Användes samma metoder på alla deltagare för att gruppera till interventions- respektive kontrollgrupp?

1.10 Är jämförelsegruppen eller -grupperna adekvat valda?

- Ja (= 3)
- Nej (= 0)

Kommentar: Finns en klar definition av jämförelsegruppen? Är det en jämförelse mellan intervention och en alternativ intervention eller mot ingen intervention alls? I den allmänna befolkningen eller i ett begränsat urval? Är jämförelsegruppen en historisk kontroll bör man vara särskilt försiktig vid värderingen.

1.11 Är det en kliniskt relevant jämförelsegrupp?

- Ja (= 3)
- Kan ej värderas (= 1)
- Nej (= 0)

2. Utfallsmått

2.1 Mättes kritiska eller viktiga effektmått/utfall på ett adekvat sätt?

- Ja (= 3)
- Kan ej värderas (= 1)
- Nej (= 0)

2.2 Var de som bedömde utfallen/resultaten medvetna om vilken intervention individerna/patienterna fått eller de undersökta exponeringsstatus?

- Ja (= 0)
- Troligen inte (= 1)
- Nej (= 2)

Kommentar: Om de som registrerar utfallen känner till vilka som fått en viss behandling kan detta öka risken för systematisk felregistrering.

2.3 Finns anledning att tro att det förekom betydande registrerings-/mätningsbias?

- Ja (= 0)
- Nej (= 3)

Kommentar: Om exempelvis expositionen är länkad till rökning och rökning i sin tur är förknippad med högre frekvens av besök i sjukvården och därför intensivare övervakning så kan registrerings/mätningsbias förekomma.

3. Jämförbarhet och förväxlingsfaktorer (confounders)

3.1 Har författarna identifierat alla viktiga förväxlingsfaktorer (confounders)?

- Ja (= 3)
- Ja, delvis (= 2)
- Nej (= 0)
- Kan ej värderas (= 1)

Lista viktiga möjliga confounders som ej har studerats:

.....

.....

Kommentar: Förväxlingsfaktorer (confounders) är variabler som påverkar utfallet i effektmåttet och vars distribution över grupperna samtidigt skiljer sig åt. Ålder, kön, bakomliggande sjukdomshistoria, riskfaktorer och socioekonomisk status är ofta särskilt viktiga.

3.2 Finns det information om hur förekomsten av förväxlingsfaktorer (confounders) skiljer sig åt mellan grupperna?

- Ja (= 0)
- Kan ej värderas (= 2)
- Nej (= 3)

Kommentar: Ett tips är att kolla den inledande tabellen (ofta) eller under bakgrundsdata (Baseline characteristics).

3.3 Är risken för selektions- eller indikationsbias stor?

- Ja, stor (= 0)
- Ja, delvis (= 1)
- Måttlig, svår att bedöma (= 2)
- Nej, liten (= 3)

Kommentar: Selektionsbias uppkommer när det i gruppernas sammansättning finns en eller flera inbyggda skillnader som i sig kan förklara skillnader i resultatet. Risken är särskilt stor när det gäller preventiva och symtomlindrande insatser som vissa välinsatta patientgrupper kan efterfråga. Risken är också stor när ett läkemedel eller en metod är särskilt lämplig att sättas in på högrisk- eller lågriskpatienter.

4. Bortfall

4.1 Har antalet exkluderade personer angivits med uppgift om orsaker?

- Ja (= 3)
- Nej (= 0)

4.2 Finns en redovisning av deltagare som förlorats från uppföljningen (bortfall och de som har valt att inte fortsätta att delta i studien)?

- Ja (= 2)
- Nej (= 0)

4.3 Hur stor är risken att bortfallet kan ha snedvridit resultaten?

- Liten (exempelvis bortfall <10%) (= 3)
- Medel (exempelvis bortfall 10–19%) (= 2)
- Stor (exempelvis bortfall 20–29%) (= 0)
- Mycket stor (exempelvis bortfall $\geq 30\%$) → undersökningen bedöms ofta sakna informationsvärde vilket kan innebära exklusion från evidensbedömning

Kommentar: Generellt ökar stora bortfall risken för att resultaten kan vara påverkade av systematiska fel. Det kan dock finnas tillfällen då man bedömer att även ett större bortfall sannolikt är slumpmässigt. Procentsiffrorna får därför ses som indikativa på om risken är stor eller liten. Bortfallet varierar också mellan olika tidpunkter i en studie och kan dessutom variera mellan olika effektmått. Bortfallet är ofta större ju längre tid som gått och därmed kan kanske data för de sista besöken vara av tveksam validitet, medan resultaten för de första besöken kan vara giltiga.

5. Undersökningens statistiska styrka (power)

5.1 Är överväganden och beräkningar till grund för val av urvalsstorlek ("sample size") tydligt beskrivna?

- Ja (= 2)
- Nej (= 0)

5.2 Hur stor blev styrkan, dvs sannolikheten för att en kliniskt relevant effekt skulle kunna påvisas med acceptabel signifikansnivå?

- Styrkan ej angiven eftersom det fanns en starkt statistiskt signifikant effekt (= 3)
- >90% (= 3)
- 80–89% (= 2)
- 70–79% (= 1)
- <70% (= 0)
- Ej angiven trots icke-signifikant resultat (= 0)

Kommentar: Små studier som inte har planerats för att kunna belysa den frågeställning som formulerats har ofta stora kvalitetsbrister i olika avseenden. Små studier har för det första inte möjlighet att besvara frågeställningen, men det är också ett tecken på att planeringen av studien varit bristfällig. Viktigt att bedöma studiens statistiska styrka (power) utifrån varje separat utfallsmått. Många studier kan t ex ha planerats för att belysa de positiva effekterna av en intervention, men inte biverkningar och risker.

6. Analys

6.1 Har den statistiska analysen av osäkerhet hanterats adekvat?

- Ja (= 3)
- Nej (= 0)

Kommentar: Redovisas konfidensintervall m m på ett adekvat sätt eller finns det en motivering för att man inte gjort det, t ex vid totalundersökningar av stora datamaterial.

6.2 Gjordes vid analysen försök att statistiskt korrigera för obalanser mellan grupperna med avseende på förväxlingsfaktorer (confounders)?

- Nej, trots att behov finns (= 0)
- Ja, genom antingen
 - matchning/restriktion (= 2)
 - stratifierad analys (= 2)
 - multivariat modellanalys, t ex regressionsanalys (= 2)
 - "propensity score"-metodik (= 2)
- Obehövt, inga obalanser av vikt bedöms finnas (= 2)

6.3 Sammanfattande bedömning: Är resultaten trovärdiga?

- Ja (= 3)
- Nej (= 0)

Kommentar: Är slutsatserna biologiskt trovärdiga? Föreligger en rimlig tidsrelation? Är sambandet starkt (mycket höga eller låga relativa risker)? Finns ett dos-respons-samband? Konsistens inom subgrupper i studien? Stämmer resultaten med fynd från andra slags forskning som randomiserade försök eller mekanistisk forskning?