

Granskningsmall för randomiserad kontrollerad prövning – modifierad från tidigare SBU-mallar (2014 och föregående)

Författare, år:

Anvisningar:

- Alternativet ”oklart” används när uppgiften inte går att få fram från texten
- Alternativet ”ej tillämpligt” väljs när frågan inte är relevant.
- Det finns förtydligande kommentarer till vissa delfrågor. Dessa anges med en fotnot.

	Ja	Nej	Oklart	Ej till- lämpligt
Extern validitet - (Directness) – Studiens överförbarhet				
1. Studiepopulationen (P i PICO, patient)				
a) Är settingen i studien relevant för svenska förhållanden idag?	☐	☐	☐	☐
Kommentar:				
b) Är den population som deltagarna togs från tydligt beskriven och relevant?	☐	☐	☐	☐
Kommentar:				
c) Beskrivs sättet att rekrytera deltagare och är det acceptabelt?	☐	☐	☐	☐
Kommentar:				
d) Stämmer studiepopulationen med HTA-rapportens frågeställning, det vill säga är inklusion- och exklusionskriterier relevanta för den specifika fråga som ställts?	☐	☐	☐	☐
Kommentar:				
e) Omfattar studiepopulationen hela P i PICO?	☐	☐	☐	☐
Kommentar:				
f) Är non-responders ¹ inkluderade i studiepopulationen?	☐	☐	☐	☐
Kommentar:				
g) Framgår det hur många personer som screenades och anges konkreta skäl för dem som inte kom att inkluderas och randomiseras?	☐	☐	☐	☐
Kommentar:				
2. Interventionen (I i PICO, intervention)				
Överensstämmer studiens intervention helt med den intervention som HTA-rapporten avser?	☐	☐	☐	☐
Kommentar:				
3. Kontroll (C i PICO, comparison)				
Överensstämmer studiens åtgärder i kontrollgruppen med de åtgärder som avses för kontroller i HTA-projektet?	☐	☐	☐	☐
Kommentar:				
4. Utfallsmått (O i PICO, outcomes)				
Överensstämmer definitionen av utfallsmått som används i studien med det/dem som HTA-rapporten avser?	☐	☐	☐	☐
Kommentar:				

	Ja	Nej	Oklart	Ej till- lämpligt
Sammanfattande bedömning av brister i överförbarhet: Inga ☐ (+ No or minor problems) Vissa ☐ (? Some problems) Stora ☐ (- Major problems)				
Motivering:				
	Ja	Nej	Oklart	Ej till- lämpligt
Intern validitet (Risk of bias – Study limitations)				
1. Selektionsbias				
a) Användes en lämplig randomiseringsmetod?	☐	☐	☐	☐
Kommentar:				
b) Är grupptilldelningen slumpmässig (random sequence generation)?	☐	☐	☐	☐
Kommentar:				
c) Är den kommande grupptillhörigheten dold fram till randomisering (concealed allocation)? ¹	☐	☐	☐	☐
Kommentar:				
d) Om studien har använt någon form av begränsning i randomiseringsprocessen (t ex block, strata), är det gjort på ett adekvat sätt? ²	☐	☐	☐	☐
Kommentar:				
e) Var grupperna sammansatta på ett tillräckligt likartat sätt?	☐	☐	☐	☐
Kommentar:				
f) Om man har korrigerat för obalanser i baslinjevariabler, har det skett på ett adekvat sätt?	☐	☐	☐	☐
Kommentar:				
Bedömning av risk för selektionsbias:	Till RevMans färgplot Låg/medelhög/hög			
Låg ☐ (+ No or minor problems)	random sequence generation			
Medelhög ☐ (? Some problems)				
Hög ☐ (- Major problems)	concealed allocation			
2. Behandlingsbias (performance bias)				
a) Var studiedeltagarna blindade? ³	☐	☐	☐	☐
Kommentar:				
b) Var behandlare/prövare blindade? ³	☐	☐	☐	☐
Kommentar:				
c) Var följsamhet i grupperna acceptabel enligt tillförlitlig dokumentation? ⁴	☐	☐	☐	☐
Kommentar:				
g) Har deltagarna i övrigt behandlats/exponerats på samma sätt bortsett från interventionen?	☐	☐	☐	☐
Kommentar:				

	Ja	Nej	Oklart	Ej till-lämpligt
Bedömning av risk för behandlingsbias: Låg ☐ (+ No or minor problems) Medelhög ☐ (? Some problems) Hög ☐ (- Major problems)				
3. Bedömningsbias (detection bias) per utfallsmått				
a) Var utfallsmåttet okänsligt för bedömningsbias?	☐	☐	☐	☐
Kommentar:				
b) Var de personer som utvärderade resultaten blindade för vilken intervention som gavs?	☐	☐	☐	☐
Kommentar:				
c) Var personerna som utvärderade utfallet opartiska?	☐	☐	☐	☐
Kommentar:				
d) Var utfallet definierat på ett lämpligt sätt?	☐	☐	☐	☐
Kommentar:				
e) Var utfallet identifierat/diagnostiserat med validerade mätmetoder?	☐	☐	☐	☐
Kommentar:				
f) Har utfallet mätts vid optimala tidpunkter?	☐	☐	☐	☐
Kommentar:				
g) Var valet av statistiskt mått för rapporterat utfall lämpligt? ⁵	☐	☐	☐	☐
Kommentar:				
h) Var den analyserade populationen (ITT eller PP) lämplig för den fråga som är föremål för studien? ⁶	☐	☐	☐	☐
Kommentar: Bedömning av risk för bedömningsbias: Låg ☐ (+ No or minor problems) Medelhög ☐ (? Some problems) Hög ☐ (- Major problems)				
4. Bortfallsbias (attrition bias) per utfallsmått				
a) Var bortfallet tillfredsställande lågt i förhållande till populationens storlek? ⁷	☐	☐	☐	☐
Kommentar:				
b) Var bortfallets storlek balanserad mellan grupperna?	☐	☐	☐	☐
Kommentar:				
c) Var relevanta baslinjevariabler balanserade mellan de som avbryter sitt deltagande och de som fullföljer studien?	☐	☐	☐	☐
Kommentar:				
d) Var den statistiska hanteringen av bortfallet adekvat?	☐	☐	☐	☐
Kommentar:				
e) Var orsakerna till bortfallet analyserade?	☐	☐	☐	☐

	Ja	Nej	Oklart	Ej till- lämpligt
Kommentar: Bedömning av risk för bortfallsbias: Låg ☐ (+ No or minor problems) Medelhög ☐ (? Some problems) Hög ☐ (- Major problems)				
5. Rapporteringsbias (selective reporting bias)				
a) Har studien följt ett i förväg publicerat studieprotokoll?	☐	☐	☐	☐
Kommentar: b) Angavs vilket/vilka utfallsmått som var primära respektive sekundära?				
	☐	☐	☐	☐
Kommentar: c) Redovisades alla i studieprotokollet angivna utfallsmått på ett fullständigt sätt? ⁸				
	☐	☐	☐	☐
Kommentar: d) Mättes biverkningar/komplikationer på ett systematiskt sätt?				
	☐	☐	☐	☐
Kommentar: e) Redovisades enbart utfallsmått som angivits i förväg i studieprotokollet? ⁹				
	☐	☐	☐	☐
Kommentar: f) Var tidpunkterna för analys angivna i förväg?				
	☐	☐	☐	☐
Kommentar: Bedömning av risk för rapporteringsbias: Låg ☐ (+ No or minor problems) Medelhög ☐ (? Some problems) Hög ☐ (- Major problems)				
6. Intressekonfliktbias				
a) Föreligger, baserat på författarnas angivna bindningar och jäv, låg eller obefintlig risk att studiens resultat har påverkats av intressekonflikter?	☐	☐	☐	☐
Kommentar: b) Föreligger, baserat på uppgifter om studiens finansiering, låg eller obefintlig risk att studien har påverkats av en finansiär med ekonomiskt intresse i resultatet?				
	☐	☐	☐	☐
Kommentar: c) Föreligger låg eller obefintlig risk för annan form av intressekonflikt (t ex att författarna har utvecklat interventionen)?				
	☐	☐	☐	☐
Kommentar: Bedömning av risk för intressekonfliktbias: Låg ☐ (+ No or minor problems) Medelhög ☐ (? Some problems) Hög ☐ (- Major problems)				

	Ja	Nej	Oklart	Ej till- lämpligt
--	----	-----	--------	----------------------

Sammanfattande bedömning av brister i intern validitet:

Inga ☐ (+ No or minor problems)

Vissa ☐ (? Some problems)

Stora ☐ (- Major problems)

Motivering:

	Ja	Nej	Oklart	Ej till- lämpligt
--	----	-----	--------	----------------------

Precision¹

a) Är studiens primära utfall ett av HTA-rapportens utfallsmått? ²

☐	☐	☐	☐
--------------------------	--------------------------	--------------------------	--------------------------

Kommentar:

b) Är minsta kliniskt relevanta effekten definierad på förhand och av rimlig storlek?

☐	☐	☐	☐
--------------------------	--------------------------	--------------------------	--------------------------

Kommentar:

c) Är de överväganden och beräkningar som ligger till grund för antal deltagare acceptabla? ³

☐	☐	☐	☐
--------------------------	--------------------------	--------------------------	--------------------------

Kommentar:

d) Om stickprovsberäkning finns (för att uppnå en viss styrka, power) stämmer studiens resultat överens med de antaganden som gjorts i denna beräkning?

☐	☐	☐	☐
--------------------------	--------------------------	--------------------------	--------------------------

Kommentar:

e) Om studiens primära syfte att visa att interventionen är bättre än jämförelsealternativet (superiority design), är konfidensintervallet så snävt att det finns en påtaglig klinisk effekt inom hela intervallet?

☐	☐	☐	☐
--------------------------	--------------------------	--------------------------	--------------------------

Kommentar:

f) Om studiens primära syfte att visa att interventionen inte är sämre än jämförelsealternativet (non-inferiority design), är konfidensintervallet så snävt att det inte finns någon påtaglig klinisk effekt inom hela intervallet?

☐	☐	☐	☐
--------------------------	--------------------------	--------------------------	--------------------------

Kommentar:

g) Är konfidensintervallen (KI) snäva? ⁴

☐	☐	☐	☐
--------------------------	--------------------------	--------------------------	--------------------------

Kommentar:

h) Vid dikotoma utfall, finns det tillräckligt många händelser? ⁵

☐	☐	☐	☐
--------------------------	--------------------------	--------------------------	--------------------------

Kommentar:

Sammanfattande bedömning av brister i precision:

Inga ☐ (+ No or minor problems)

Vissa ☐ (? Some problems)

Stora ☐ (- Major problems)

Motivering:

Precision för övriga utfallsmått ("O") i HTA-rapporten (ett i taget) bedöms vid behov (kopiera precisionsfrågorna).

Förklaringar

Extern validitet - (Directness) – Studiens överförbarhet

¹ Vid test av en behandling bör alla som utsatts för testet inkluderas i den kommande studiepopulationen och inte bara de som svarat positivt på behandlingen.

Intern validitet (Risk of bias – Study limitations)

¹ Med ”concealed allocation” avses att det inte ska vara möjligt för patient, behandlare eller provare att förutse till vilken grupp patienten kommer att randomiseras.

² Risk för att randomiseringen blir förutsägbar för provare eller deltagare kan t ex finnas vid blockrandomisering. Blockrandomisering används bl a i multicenterstudier för att motverka slumpmässiga snedfördelningar mellan olika centra eller länder.

³ Det är önskvärt att både patienter och provare är blindade i en studie. Ibland kan det av praktiska skäl vara svårt eller omöjligt att dölja för provare och/eller patient vilken behandling som ges. Däremot är det oftast möjligt att den som utvärderar effekten är blindad.

De alternativ som finns är följande:

- Öppen prövning: ingen är blindad
- Enkelblind: a) deltagarna är blindade; b) provare och/eller resultatutvärderare är blindade
- Dubbelblind: a) deltagarna och provare och/eller resultatutvärderare är blindade; b) deltagarna och provare och/eller resultatutvärderare är blindade samt det framgår av beskrivningen att all resultatbearbetning utfördes innan provningskoden bröts.

Det finns flera exempel på studier där blindningen inte har lyckats pga karakteristiska effekter eller biverkningar av aktiv behandling, exempelvis muntorrhet vid behandling med neuroleptika och underlivsblödningar vid behandling med östrogen. I vissa fall är det möjligt att ge biverkningsmotverkande medel som tillägg till aktiv behandling för att minska risken för att blindningen äventyras. Andra faktorer som kan försvåra blindningen är bristande likhet mellan tabletter, inhalationspreparat etc avseende utseende eller smak.

En stor ”placeboeffekt” i kontrollgruppen kan tala för en lyckad blindning. I vissa studier låter man studiedeltagarna gissa om de fått aktiv behandling eller kontroll.

⁴ Kontroll av följsamheten är särskilt viktig då man inte ser en signifikant effektskillnad mellan grupperna. En bristande följsamhet kan minska såväl interventionens effekter som bieffekter. Om interventionen har en signifikant effekt är kontroll av följsamheten ofta av mindre betydelse. Undantag är om det var sämre följsamhet i gruppen som fick referensbehandling. Det senare är tänkbart i en placebokontrollerad studie om blindningen varit otillräcklig, alternativt om en referensbehandling har mycket högre frekvens av biverkningar.

Ett riktmärke för vad som anses vara acceptabel följsamhet är att mer än 80 procent av deltagarna ska ha mottagit mer än 80 procent av behandlingen.

⁵ De vanligaste mått som används för dikotoma variabler är riskkvot (”risk ratio”, RR), oddskvot (”odds ratio”, OR), absolut riskreduktion/riskskillnad (”risk difference”) och ”number needed to treat” (NNT).

För kontinuerliga variabler används vanligen absolut skillnad i medelvärde (”difference in means”, ”mean difference”). Alla måtten ska redovisas med lämpligt spridningsmått, företrädesvis 95 procents konfidensintervall.

⁶ Resultaten kan analyseras enligt ”intention to treat” (ITT) och/eller per protokoll (PP). En ITT-analys innebär att alla personer som randomiserats följs upp inom sin behandlingsarm oavsett om de fått tilldelad behandling eller inte och är oftast den metod som bör användas. Om resultaten är beräknade på annat sätt än med ITT finns det risk för att behandlingseffekten blir överskattad. ITT-analysen kan

kompletteras med en känslighetsanalys enligt ”worst case scenario” där sämsta tänkbara utfall tillskrivs saknade patienter i den grupp som uppvisar bäst effekt och bästa tänkbara utfall tillskrivs saknade patienter i den grupp som uppvisar sämst effekt. Ibland är det önskvärt att även en PP-analys redovisas, vilket innebär att bara de som följt hela studieprotokollet ingår i analysen.

Vid bortfall i studier där man använder kontinuerliga variabler eller skalor använder man ibland en beräkningsmetod där det senast uppmätta resultatet antas gälla även för senare tidpunkter då data saknas (”last observation carried forward”, LOCF).

⁷ Det bortfall som bedöms här avser bortfall efter randomisering. Det kan dock finnas tillfällen då även ett större bortfall sannolikt är slumpmässigt. Nedanstående exempel bör därför ses som riktvärden.

- Litet (<10%)
- Måttligt (10–19%)
- Stort (20–29%)
- Mycket stort ($\geq 30\%$). Undersökningen bedöms ofta sakna informationsvärde vilket kan innebära att studien bör exkluderas.

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

⁸ Även om redovisade effektmått är rimliga, definierade i förväg och adekvat rapporterade kan det finnas andra viktiga effektmått som utelämnats.

⁹ Det är inte ovanligt att studier med negativa resultat inkluderar förklarande efteranalyser (”explanatory”- eller ”post hoc”-analyser) för att t ex finna vissa subgrupper inom den studerade patientgruppen som kan ha nytta av behandlingen. Dessa analyser kan fylla en viktig hypotes-genererande funktion, men är naturligtvis behäftade med stor osäkerhet. Slutsatserna i en studie får därför aldrig baseras på sådana analyser.

Precision

¹ Frågorna a-f gäller studiens primära utfallsmått. Om studiens primära utfallsmått ingår bland HTA-projektets utfallsmått bör det finnas en stickprovsberäkning och svaren på frågorna kan ingå i bedömningen av just det utfallets precision.

För övriga utfall bedöms precision endast enligt frågorna e-g.

² Om studiens primära utfallsmått inte ingår bland HTA-projektets utfallsmått, blir frågorna b-e irrelevanta för HTA-projektet, men kan ge en allmän bild av studiens kvalitet.

³ Beräkning av stickprovsstorlek, ibland kallad ”power”-beräkning, används för att beräkna studiers statistiska styrka, dvs för att räkna ut hur stor en studie bör vara för att med rimlig sannolikhet kunna påvisa en behandlingseffekt. Det är viktigt att författarna redovisar hur de kommit fram till vald studiestorlek och att beräkningen skett före studiestart. Annars kan det inte uteslutas att författarna successivt adderat deltagare till studien tills statistisk signifikans har uppnåtts.

⁴ Vid påvisad skillnad mellan grupperna bör hela KI omfatta en kliniskt relevant effekt, dvs den nedre gränsen bör ligga vid eller ovan den minsta kliniskt relevanta effekten. Vid ingen påvisad skillnad mellan grupperna bör KI vara så pass snäva att det inte finns påtaglig förväntad klinisk effekt inom intervallet.

⁵ Minst 200 händelser har använts som en tumregel i Cochrane handbook.