



CHECKLISTA FÖR RANDOMISERADE KONTROLLERADE STUDIER

Verktåg för kritisk bedömning att användas i JBI:s systematiska översikter

INTRODUKTION

JBI är en internationell forskningsorganisation baserad vid fakulteten för hälsovetenskaper och medicinska vetenskaper vid University of Adelaide, South Australia. JBI utvecklar och levererar unik evidensbaserad information, programvara och allmän och yrkesinriktad utbildning utformat för att förbättra hälso- och sjukvårdspraxis och hälsoresultat. Med över 70 samarbetande enheter som betjänar över 90 länder är JBI en erkänd global ledare inom evidensbaserad hälso- och sjukvård.

JBI:s systematiska litteraturöversikter

Kärnan i evidenssynes är den systematiska genomgången av litteratur för en viss intervention, ett visst förhållande eller ett visst problem. Den systematiska genomgången är i huvudsak en analys av tillgänglig litteratur (det vill säga evidens) och en bedömning av den eventuella effektiviteten av en praxis, vilket inbegriper en rad komplexa steg. JBI intar en viss hållning till vad som räknas som evidens och vilka metoder som ska användas för att syntetisera dessa olika typer av evidens. I enlighet med denna bredare syn på evidens har JBI utvecklat teorier, metoder och rigorösa processer för kritisk bedömning och syntetisering av dessa olika former av evidens för att underlätta kliniskt beslutsfattande inom hälso- och sjukvården. Det finns nu JBI-vägledning för att genomföra granskningar av effektivitetsforskning, kvalitativ forskning, prevalens/incidens, etiologi/risk, hälsoekonomiska utvärderingar, textartiklar/sakkunnigutlåtanden, diagnostisk noggrannhet hos test, mixade metoder, umbrella reviews och scoping reviews. Mer information om JBI:s systematiska litteraturöversikter finns i [JBI:s evidenssynteshandbok](#).

JBI:s verktyg för kritisk bedömning

Alla systematiska litteraturöversikter omfattar en process för kritik eller bedömning av forskningsevidensen. Syftet med denna bedömning är att utvärdera en studies metodologiska kvalitet och att fastställa i vilken utsträckning en studie har tagit upp möjligheten till bias i dess utformning samt vid dess utförande och analys. Alla artiklar som väljs ut för att inkluderas i den systematiska litteraturöversikten (det vill säga, de artiklar som uppfyller inklusionskriterierna som beskrivs i protokollet) måste bli föremål för noggrann bedömning av två kritiska bedömare. Resultaten av denna bedömning kan sedan användas för att informera syntes och tolkning av resultaten av studien. JBI:s verktyg för kritisk bedömning har utvecklats av JBI och samarbetspartner och godkänts av JBI:s vetenskapliga kommitté efter omfattande peer review. Även om JBI:s verktyg för kritisk bedömning är utformade för att användas i systematiska litteraturöversikter kan de också användas när ni skapar Critically Appraised Topics (CAT), i tidskriftsklubbar och som ett pedagogiskt verktyg.

JBI:S CHECKLISTA FÖR KRITISK BEDÖMNING AV RANDOMISERADE KONTROLLERADE STUDIER

Bedömare _____ Datum _____

Författare _____ År _____ Journalnummer _____

	Ja	Nej	Oklart	Ej tillämpligt
1. Användes äkta randomisering för tilldelning av deltagare till behandlingsgrupper?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Doldes allokeringen till grupper?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Var behandlingsgrupperna likartade vid baslinjen?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Var deltagarna blindade för behandlingstilldelningen?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Var de som gav behandling blindade för behandlingstilldelningen?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Var utfallsbedömarna blindade för behandlingstilldelningen?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Behandlades behandlingsgrupperna på samma sätt bortsett från interventionen av intresse?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Var uppföljningen fullständig, och om inte, var skillnader mellan grupper vad beträffar uppföljning av dem beskrivna och analyserade på ett tillfredsställande sätt?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Analyserades deltagarna i de grupper som de randomiserades till?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Mättes utfallen på samma sätt för behandlingsgrupper?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. Mättes utfallen på ett tillförlitligt sätt?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. Användes lämplig statistisk analys?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. Var försöksdesignen lämplig, och beaktades eventuella avvikelser från standarddesignen för randomiserade kontrollerade studier (individuell randomisering, parallella grupper) vid genomförandet och analysen av studien?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Sammanfattande bedömning: Inkludera Exkludera Sök ytterligare information

Kommentarer (inklusive skäl för exklusion)

FÖRKLARING TILL VERKTYGET FÖR KRITISK BEDÖMNING AV RANDOMISERADE KONTROLLERADE STUDIER MED ENSKILDA DELTAGARE I PARALLELLA GRUPPER

Så här citerar du: Tufanaru C, Munn Z, Aromataris E, Campbell J och Hopp, L. 2020. Chapter 3: Systematic reviews of etiology and risk. I: Aromataris, E. och Munn, Z. (red.). *JBI Manual for Evidence Synthesis*. JBI. Kan läsas på eller laddas ned från <https://synthesismanual.jbi.global>

Svar: Ja, Nej, Oklart eller Ej tillämpligt

Verktuget för kritisk bedömning av randomiserade kontrollerade studier (enskilda deltagare i parallella grupper)

1. Användes äkta randomisering för tilldelning av deltagare till behandlingsgrupper?

Skillnaderna mellan deltagarna i jämförda grupper utgör ett hot mot den interna validiteten i en studie som undersöker orsakssamband. Om deltagarna inte allokeras till behandlings- och kontrollgrupper genom slumpmässig tilldelning finns det en risk för att allokeringen påverkas av deltagarnas kända karakteristika, och dessa skillnader mellan grupperna kan då snedvrída gruppernas jämförbarhet. En äkta slumpmässig tilldelning av deltagare till grupperna innebär att en procedur används för allokering av deltagarna till grupper enbart baserat på slumpen utan inverkan av deltagarnas kända karakteristika. Kontrollera detaljerna om randomiseringsproceduren som användes för allokering av deltagarna till studiegrupper. Användes en procedur med äkta slumpmässighet (äkta randomisering)? Användes till exempel en lista med slumpstal? Användes en datorgenererad lista med slumpstal?

2. Doldes allokeringen till grupper?

Om de som allokerar deltagare till de jämförda grupperna vet vilken grupp som står på tur i allokeringsprocessen, det vill säga, behandlings- eller kontrollgruppen, finns det risk för att de avsiktligt och för egna ändamål ingriper i allokeringen av patienter genom att företrädesvis allokera patienter till antingen behandlingsgruppen eller till kontrollgruppen. Detta kan då snedvrída genomförandet av den allokeringsprocess som anges i randomiseringen, och resultaten av studien kan därför vara förvrängda i sådant fall. Med döljande av allokering avses förfaranden som hindrar dem som allokerar patienter från att före allokeringen görs veta vilken behandlings- eller kontrollgrupp som står på tur i allokeringsprocessen. Kontrollera detaljerna om förfarandet som användes för döljande av allokering. Användes ett lämpligt förfarande för döljande av allokering? Användes till exempel central randomisering? Användes sekventiellt numrerade, ogenomskinliga och förseglade kuvert? Användes kodade läkemedelsförpackningar?

3. Var behandlingsgrupperna likartade vid baslinjen?

Skillnaderna mellan deltagarna i jämförda grupper utgör ett hot mot den interna validiteten i en studie som undersöker orsakssamband. Om det finns skillnader mellan deltagare i jämförda grupper finns det risk för selektionsbias. Om det finns skillnader mellan deltagare i de jämförda grupperna är det inte säkert att "effekten" kan

tillskrivas den potentiella "orsaken" (den undersökta interventionen eller behandlingen) eftersom det är rimligare att "effekten" kan förklaras av skillnaderna mellan deltagarna, det vill säga av en selektionsbias. Kontrollera de karakteristika som rapporterats för deltagarna. Är deltagarna från de jämförda grupperna likartade i fråga om de karakteristika som potentiellt kan förklara effekten, även vid brist på "orsaken", till exempel ålder, sjukdomens svårighetsgrad, stadium av sjukdomen, samtidiga tillstånd, och så vidare? Kontrollera andelen deltagare med specifika relevanta karakteristika i de jämförda grupperna. Kontrollera medelvärdena för relevanta mätningar i de jämförda grupperna (smärtpoäng, ångestpoäng, osv.). *[Obs! Beakta INTE bara p-värdet för statistisk testning av skillnaderna mellan grupperna i fråga om baslinjekarakteristika.]*

4. Var deltagarna blindade för behandlingstilldelningen?

Om deltagarna vet om att de allokerats till behandlingsgruppen eller till kontrollgruppen finns det risk för att de uppträder annorlunda och svarar eller reagerar på ett annat sätt på interventionen av intresse respektive kontrollinterventionen jämfört med om de inte vet om detta, och resultaten av studien kan därför vara förvrängda i sådant fall. Blindning av deltagare används för att minimera denna risk. Blindning av deltagarna avser förfaranden som hindrar deltagarna från att veta vilken grupp de allokeras till. Om blindning av deltagarna används vet inte deltagarna om de är i gruppen som får behandlingen av intresse eller om de är i någon annan grupp som får kontrollinterventionerna. Kontrollera de uppgifter som rapporteras i artikeln om blindning av deltagare i fråga om behandlingstilldelning. Användes ett lämpligt blindningsförfarande? Användes till exempel identiska kapslar eller sprutor? Användes identiska anordningar? Var uppmärksam på användning av olika termer för samma sak; blindning kallas ibland även maskning.

5. Var de som gav behandling blindade för behandlingstilldelningen?

Om de som ger behandling vet om att deltagare allokerats till behandlingsgruppen eller till kontrollgruppen finns det risk för att de uppträder annorlunda mot deltagarna från behandlingsgruppen i förhållande till deltagarna från kontrollgruppen, eller risk för att de ger behandling på ett annorlunda sätt, jämfört med om de inte vet om detta. I sådant fall kan detta inverka på genomförandet av de jämförda behandlingarna, och resultaten av studien kan därför vara förvrängda. Blindning av de som ger behandling används för att minimera denna risk. Blindning av de som ger behandling avser förfaranden som hindrar de som ger behandling från att veta vilken grupp de behandlar. Med andra ord vet inte de som ger behandling om de behandlar gruppen som får behandlingen av intresse eller om de behandlar

någon annan grupp som får kontrollinterventionerna. Kontrollera de uppgifter som rapporteras i artikeln om blindning av de som ger behandling i fråga om behandlingstilldelning. Finns det någon information i artikeln om de som ger behandling? Var de som gav behandling ovetande om vilken av de jämförda grupperna som respektive tilldelning av deltagare gjordes till?

6. Var utfallsbedömarna blindade för behandlingstilldelningen?

Om de som bedömer utfallen vet om att deltagare allokerats till behandlingsgruppen eller till kontrollgruppen finns det risk för att de uppträder annorlunda mot deltagarna från behandlingsgruppen i förhållande till deltagarna från kontrollgruppen jämfört med om de inte vet om detta. Då finns det risk för att mätningen av utfallen är snedvriden, och resultaten av studien kan därför vara förvrängda. Blindning av utfallsbedömare används för att minimera denna risk. Kontrollera de uppgifter som rapporteras i artikeln om blindning av utfallsbedömare i fråga om behandlingstilldelning. Finns det någon information i artikeln om utfallsbedömare? Var de som bedömde behandlingens effekter på utfall ovetande om vilken av de jämförda grupperna som respektive tilldelning av deltagare gjordes till?

7. Behandlades behandlingsgrupperna på samma sätt bortsett från interventionen av intresse?

För att "effekten" ska kunna tillskrivas "orsaken" (behandlingen eller interventionen av intresse), förutsatt att det inte finns någon selektionsbias, får det inte finnas någon annan skillnad mellan grupperna vad beträffar behandling eller vård som erhållits förutom den manipulerade "orsaken" (den behandling eller intervention som kontrolleras av forskarna). Om det förekommer andra exponeringar eller behandlingar samtidigt med "orsaken" (behandlingen eller interventionen av intresse), förutom "orsaken", kan eventuellt inte "effekten" tillskrivas den undersökta "orsaken" (den undersökta behandlingen) eftersom det är rimligare att "effekten" kan förklaras av andra exponeringar eller behandlingar som förekommer samtidigt med "orsaken" (behandlingen av intresse). Kontrollera de rapporterade exponeringar eller interventioner som de jämförda grupperna erhållit. Finns det andra exponeringar eller behandlingar som förekommer samtidigt med "orsaken"? Är det rimligare att "effekten" kan förklaras av andra exponeringar eller behandlingar som förekommer samtidigt med "orsaken"? Är det helt säkert att det inte finns någon annan skillnad mellan grupperna vad beträffar behandling eller vård som erhållits förutom behandlingen eller interventionen av intresse?

8. Var uppföljningen fullständig, och om inte, var skillnader mellan grupper vad beträffar uppföljning av dem beskrivna och analyserade på ett tillfredsställande sätt?

För den här frågan avser "uppföljning" tidsperioden från tidpunkten för randomiserad allokering (randomiserad tilldelning, eller randomisering) till jämförda grupper fram tills tidpunkten för avslutande av studien. Denna fråga vid kritisk bedömning söker svar på om det finns fullständig kunskap (mätningar, observationer osv.) för studiens hela varaktighet enligt

vad som definierades i föregående mening (det vill säga, från tidpunkten för randomiserad allokering fram tills tidpunkten för avslutande av studien) om alla slumpmässigt allokerade deltagare. Om det förekommer ofullständig uppföljning, det vill säga ofullständig kunskap om alla slumpmässigt allokerade deltagare, kallas detta i metodlitteraturen för bortfall efter randomisering. Eftersom randomiserade kontrollerade studier inte är perfekta förekommer det nästan alltid bortfall efter randomisering, och fokuset för denna fråga ligger på lämplig undersökning av bortfall efter randomisering (beskrivning av utebliven uppföljning, beskrivning av orsakerna till utebliven uppföljning, uppskattning av inverkan av utebliven

uppföljning på effekterna, osv.). Om det finns skillnader i fråga om utebliven uppföljning mellan de jämförda grupperna i en randomiserad kontrollerad studie utgör dessa skillnader ett hot mot den interna validiteten i en randomiserad experimentell studie som undersöker kausala effekter eftersom dessa skillnader kan ge en rimlig alternativ förklaring till den observerade "effekten", även vid brist på "orsaken" (behandlingen eller interventionen av intresse). Vid kritisk bedömning av en randomiserad kontrollerad studie, kontrollera om det fanns skillnader i fråga om utebliven uppföljning mellan de jämförda grupperna. Om uppföljningen var ofullständig (det vill säga att informationen var ofullständig om alla deltagare), undersök de rapporterade detaljerna om de strategier som användes för att ta itu med ofullständig uppföljning, t.ex. beskrivningar av utebliven uppföljning (absoluta tal, proportioner, orsaker till utebliven uppföljning) och granska konsekvensanalyser (analyserna av inverkan av utebliven uppföljning på resultaten). Fanns det en beskrivning av den ofullständiga uppföljningen (antal deltagare och de specifika orsakerna till utebliven uppföljning)? Det är viktigt att notera att, i fråga om utebliven uppföljning, det inte räcker med att bara veta antalet deltagare och andelen deltagare med ofullständiga uppgifter. Orsakerna till utebliven uppföljning är väsentliga vid analysen av risk för bias. Även om antalet deltagare och andelen deltagare med ofullständiga uppgifter är likartad eller identisk i jämförda grupper kan det medföra risk för bias om mönstren i skälen till utebliven uppföljning skiljer sig åt (t.ex. biverkningar orsakade av interventionen av intresse, utebliven kontakt, osv.) och dessa mönster inte utforskas och beaktas på lämpligt sätt i analysen. Om det finns skillnader mellan grupperna i fråga om utebliven uppföljning (siffror/proportioner och orsaker), fanns det då en analys av mönstren till utebliven uppföljning? Om det finns skillnader mellan grupperna i fråga om utebliven uppföljning, fanns det då en analys av inverkan av utebliven uppföljning på resultaten? [Obs! Fråga 8 handlar INTE om en intention to treat-analys, vilket är föremålet för fråga 9.]

9. Analyserades deltagarna i de grupper som de randomiserades till?

Denna fråga handlar om ITT-analysen (intention to treat-analysen). Det finns olika statistiska analysstrategier att tillgå för analys av data från randomiserade kontrollerade studier, t.ex. intention to treat-analys (även kallat intent to treat-analys, förkortat ITT-analys), per protokoll-analys och as treated-analys. I ITT-analysen analyseras deltagarna i de grupper som de randomiserades till, oavsett om de faktiskt deltog eller inte i dessa grupper under studiens hela varaktighet, erhöll den experimentella interventionen eller kontrollinterventionen som

planerat eller inte, eller var följsamma eller inte med den planerade experimentella interventionen eller kontrollinterventionen. I ITT-analysen jämförs utfallen för deltagarna från de ursprungliga grupper som skapades genom den första randomiserade allokeringen av deltagare till de grupperna. Kontrollera om ITT-analysen rapporterades, och kontrollera detaljerna i ITT-analysen. Analyserades deltagarna i de grupper som de först randomiserades till, oavsett om de faktiskt deltog eller inte i dessa grupper, och oavsett om de faktiskt erhöll de planerade interventionerna eller inte? *[Obs! ITT-analysen är en typ av statistisk analys som rekommenderas i Consolidated Standards of Reporting Trials (CONSORT) om bästa praxis vid rapportering av studier, och den anses vara en markör för god metodologisk kvalitet i analysen av resultaten från en randomiserad studie. ITT uppskattar effekten av att erbjuda interventionen, det vill säga effekten av att instruera deltagarna att använda eller erhålla interventionen. ITT uppskattar inte effekten av att faktiskt erhålla interventionen av intresse.]*

10. Mättes utfallen på samma sätt för behandlingsgrupper?

Om utfallet ("effekten") inte mäts på samma sätt i de jämförda grupperna finns det ett hot mot den interna validiteten i en studie som undersöker ett orsakssamband eftersom skillnaderna i utfallsmätningar kan förväxlas med en effekt av behandlingen ("orsaken"). Kontrollera om utfallen mättes på samma sätt. Användes samma instrument eller skala? Var det samma mättid? Var metoderna och instruktionerna för mätning samma?

11. Mättes utfallen på ett tillförlitligt sätt?

Bristande reliabilitet hos utfallsmätningar är ett hot som försvagar validiteten för inferenser om det statistiska förhållandet mellan "orsaken" och "effekten" som uppskattats i en studie som undersöker kausala effekter. Bristande reliabilitet hos utfallsmätningar är en av de olika rimliga förklaringarna till fel i den statistiska inferensen i fråga om förekomst och omfattning av den effekt som bestäms av behandlingen ("orsaken"). Kontrollera detaljerna om mätningens reliabilitet, t.ex. antal bedömare, bedömarnas utbildning, intrabedömarreliabilitet och interbedömarreliabilitet i studien (inte som rapporterats i externa källor). Denna fråga handlar om reliabiliteten i den mätning som utförts i studien; det handlar inte om validiteten hos de mätinstrument/skalor som använts i studien. *[Obs! Två andra viktiga hot som försvagar validiteten i inferenser om det statistiska förhållandet mellan "orsaken" och "effekten" är låg statistisk styrka och överträdelser av antaganden i statistiska test. Dessa andra två hot undersöks i fråga 12).]*

12. Användes lämplig statistisk analys?

Olämplig statistisk analys kan orsaka fel i den statistiska inferensen i fråga om förekomst och omfattning av den effekt som bestäms av behandlingen ("orsaken"). Låg statistisk styrka och överträdelser av antaganden i statistiska test är två viktiga hot som försvagar validiteten av inferenser om det statistiska förhållandet mellan "orsaken" och "effekten". Kontrollera följande aspekter: Huruvida antagandena i statistiska test respekterades; huruvida lämplig analys av statistisk styrka utfördes; huruvida lämpliga effektstorlekar användes; huruvida

lämpliga statistiska förfaranden eller metoder användes med tanke på antalet och typen av beroende och oberoende variabler, antalet studiegrupper, beskaffenheten av förhållandet mellan grupperna (oberoende eller beroende grupper) och målen med statistisk analys (sambandet mellan variabler, förutsägelse, överlevnadsanalys osv.).

13. Var försöksdesignen lämplig för ämnet, och beaktades eventuella avvikelser från standarddesignen för randomiserade kontrollerade studier vid genomförandet och analysen av studien?

Randomiserade kontrollerade studier med vissa designer, t.ex. randomiserade kontrollerade studier med mixad metod (crossover RCT-studier), ska endast genomföras när så är lämpligt. Alternativa designer kan också innebära ytterligare risker för bias om de inte beaktas vid utformningen och analysen.

Studier med mixad metod (crossover-studier) ska endast utföras på personer med kroniskt, stabilt tillstånd där interventionen ger en kortvarig effekt (dvs. lindring av symtom). Crossover-studier ska säkerställa att det finns en lämplig washout-period mellan behandlingarna.

Randomiserade kontrollerade studier med klusterrandomisering randomiserar grupper av individer så att de bildar "kluster". När vi bedömer utfall på en individuell nivå i prövningar som genomförs i kluster finns det problem med analysenheter eftersom individer inom ett kluster är korrelerade med varandra. Detta ska beaktas av studieförfattarna när de utför analys, och helst ska författarna ha rapporterat om korrelationskoefficienten inom klustret.

Stepped wedge randomiserade kontrollerade studier kan vara lämpliga när det förväntas att interventionen kommer att göra mer nytta än skada, eller på grund av logistiska, praktiska eller ekonomiska faktorer vid utrollningen av en ny behandling/intervention. Dataanalys vid dessa prövningar ska utföras på lämpligt sätt med beaktande av effekter av tid.