

JBI: Kriittisen arvioinnin tarkistuslista satunnaistetulle kontrolloidulle tutkimukselle (RCT)

Tätä tarkistuslistaa käytetään satunnaistetun kontrolloidun tutkimuksen (randomized controlled trial, RCT) metodologisen laadun arviointiin ja tutkimuksen tuloksiin vaikuttavan harhan riskin tunnistamiseen. Arvioinnin tarkistuslistaan sisältyy yhteensä 13 arviointikriteeriä, joiden yksityiskohtaiset sisällöt on kuvattu alla. Arvioijan on hyvä tutustua myös JBI:n julkaisemaan katsauksen tekijöiden käsikirjaan, jonka pdf-linkki löytyy Tutkimusten arviointikriteeristö -kotisivulta arviointia tehdessään. Tarkistuslistan alkuperäinen englanninkielinen versio löytyy kotisivumme kohdasta alkuperäiset tutkimusten arviointikriteeristöt. Kunkin kriteerin toteutuminen arvioidaan asteikolla: Kyllä (K), Ei (E), Epäselvä (?), Ei sovellettavissa (NA). (Tufanaru ym. 2017.)

Arvioija _____ Päiväys _____

Tekijä(t) _____ Vuosi _____ Nro _____

Arviointikriteeri	K	E	?	NA
1. Onko osallistujien ryhmiin jakaminen satunnaistettu?	☐	☐	☐	☐
2. Ovatko tutkittavien ryhmiin jako salattu ryhmiin jakoa toteuttaneilta?	☐	☐	☐	☐
3. Ovatko koe- ja kontrolliryhmät samankaltaisia tutkimuksen alussa?	☐	☐	☐	☐
4. Ovatko tutkittavat sokkoutettu tutkimuksen ryhmäjaoista?	☐	☐	☐	☐
5. Ovatko intervention toteuttajat sokkoutettu tutkittavien ryhmäjaoista?	☐	☐	☐	☐
6. Ovatko tulostuottajien mittajat sokkoutettu tutkittavien ryhmäjaoista?	☐	☐	☐	☐
7. Kohdeltiinko ryhmiä yhdenmukaisesti lukuun ottamatta tutkimuksen kohteena olevaa interventiota?	☐	☐	☐	☐
8. Pysyivätkö tutkittavat mukana tutkimuksessa seurannan aikana, ja elleivät pysyneet, kuvattiinko ja analysoitiinko seurannan aikana ilmenneet ryhmien väliset erot asianmukaisesti?	☐	☐	☐	☐
9. Tehtiinkö lähtöryhmien mukainen (hoitoaieanalyysi eli 'intention-to-treat') analyysi?	☐	☐	☐	☐
10. Mitattiinko muuttujat samalla tavalla kaikissa ryhmissä?	☐	☐	☐	☐
11. Mitattiinko muuttujat luotettavasti?	☐	☐	☐	☐
12. Käytettiin soveltuvia tilastollisia menetelmiä?	☐	☐	☐	☐
13. Onko koeasetelma tutkittavan aihealueen näkökulmasta asianmukainen, ja huomioitiinko mahdolliset poikkeavuudet perinteisestä RCT-asetelmasta tutkimuksen toteutuksessa ja analyysissä?	☐	☐	☐	☐

Kokonaisarviointi: Hyväksy ☐ Hylkää ☐ Lisätietoja tarvitaan ☐

Kommentit (mukaan lukien syy hylkäykseen):

Kriittisen arvioinnin tarkistuslista

Lähde: Tufanaru C, Munn Z, Aromataris E, Campbell J, Hopp L. Kappale 3: Systematic reviews of effectiveness. Teoksessa: Aromataris E, Munn Z (toim.). Joanna Briggs Institute Reviewer's Manual. The Joanna Briggs Institute, 2017. Saatavilla <https://reviewersmanual.joannabriggs.org/>

1. Oliko osallistujien ryhmiin jakaminen satunnaistettu?

Tutkimuksen, jossa tutkitaan syy-seuraussuhteita, vertailuryhmien tulisi olla osallistujiltaan mahdollisimman samankaltaiset tutkimuksen sisäisen validiteetin vahvistamiseksi. Jos osallistujia ei jaeta satunnaisesti koe- ja kontrolliryhmiin, on vaarana, että jakamiseen vaikuttavat osallistujien tietyt tiedossa olevat ominaisuudet, ja ryhmien väliset erot voivat vääristää ryhmien vertailukelpoisuutta. Osallistujien todellinen satunnaistaminen ryhmiin tarkoittaa sitä, että hyödynnetään sellaista menettelytapaa, jossa osallistujia jaetaan ainoastaan sattumanvaraisesti eikä osallistujien ominaisuuksien perusteella.

Arvio: Tarkista annetut tiedot osallistujien ryhmäjaossa käytetystä satunnaistamisen menettelytavasta. Käytettiinkö sattumanvaraista menettelytapaa? Oliko esimerkiksi satunnaislukutaulukko käytössä? Hyödynnettiinkö tietokoneavusteista satunnaislukutaulukkoa?

2. Onko tutkittavien ryhmiin jako salattu ryhmiin jakoa toteuttaneilta?

Jos tutkimuksen osallistujia vertailuryhmiin jakavat ovat tietoisia siitä, kuuluvatko osallistujat koe- tai kontrolliryhmiin, on vaarana, että jakajat jakavat osallistujia tietoisesti ja tarkoituksenmukaisesti joko koe- tai kontrolliryhmiin. Tämä voi vääristää satunnaistetun jakoprosessin toteutumista, jonka seurauksena myös tutkimustulokset voivat olla vääristyneitä. Salaaminen viittaa menettelytapoihin, jotka estävät tutkimukseen osallistuvien ryhmäjaosta vastaavia henkilöitä tietämästä, onko jakoprosessissa meneillään koe- vai kontrolliryhmän muodostaminen.

Arvioi: Tarkista tarkemmat yksityiskohdat salaamisen menettelytavoista. Oliko asianmukainen menettelytapa käytössä? Esimerkiksi, käytettiinkö keskitettyä satunnaistamista? Oliko peräkkäisesti numeroidut, läpinäkymättömät ja suljetut kirjekuoret käytössä? Käytettiinkö koodattuja lääkepakkauksia?

3. Ovatko koe- ja kontrolliryhmät samankaltaisia tutkimuksen alussa?

Syy-seuraus-suhteita tarkastelevassa tutkimuksessa vertailtavien ryhmien osallistujien väliset erot ovat uhka tutkimuksen sisäiselle validiteetille. Jos vertailuryhmien osallistujien välillä on eroja, on mahdollista, ettei 'seuraus' ole tarkasteltavan 'syyn' (tutkimuksen kohteena oleva interventio tai hoitomenetelmä) seurausta. Tällöin on mahdollista, että 'seuraus' voidaan selittää osallistujien välisillä eroilla, toisin sanoen valintaharhalla.

Arvioi: Tarkista raportoidut tiedot osallistujien ominaisuuksista. Ovatko vertailuryhmien osallistujat samankaltaisia sellaisten ominaisuuksien osalta, jotka voivat selittää seurauksen

‘syyn’ (esimerkiksi ikä, taudin vakavuus, taudin vaihe, yhtäaikaista sairautta, jne.)? Tarkista vertailuryhmien osallistujien olennaisten ominaisuuksien suhteelliset osuudet. Tarkista vertailuryhmistä tehtyjen mittausten keskiarvot (esimerkiksi kipumittarin tai ahdistuneisuuskyselyn tulokset). [Huom: ÄLÄ ota huomioon ainoastaan tilastollisten testien P-arvoja vertailuryhmien lähtötason ominaisuuksien välisten erojen tarkastelussa.]

4. Ovatko tutkittavat sokkoutettu tutkimuksen ryhmäjäoista?

Jos osallistujat ovat tietoisia siitä, kuuluvatko he koe- vai kontrolliryhmään, on vaarana, että tutkimustulokset ovat vääristyneitä. Osallistujat voivat käyttäytyä ja reagoida eri tavalla koeryhmässä tai vastaavasti kontrolliryhmässä verrattuna siihen, jos he eivät ole tietoisia kumpaan ryhmään he kuuluvat. Tämän riskin minimoimiseksi se, kumpaan ryhmään tutkittava kuuluu, salataan osallistujilta. Jos osallistujien sokkouttamista käytetään, osallistujat eivät ole tietoisia siitä, onko heidän saamansa interventio tai hoito koeinterventio, tai ovatko he ryhmässä, joka saa kontrolli-interventiota.

Arvioi: Tarkista raportoiduista tiedoista, onko osallistujien ryhmiin jaossa käytetty sokkouttamista. Oliko asianmukainen sokkouttamisen menettelytapa käytössä? Esimerkiksi, käytettiinkö saman näköisiä kapseleita tai ruiskuja? Olivatko samanlaiset laitteet käytössä? Huomio se, että sokkouttamisesta voidaan käyttää erilaisia englanninkielisiä käsitteitä kuten ‘masking’ tai ‘blinding’.

5. Ovatko intervention toteuttajat sokkoutettu tutkittavien ryhmäjäoista?

Jos intervention tai hoitomenetelmän toteuttajat ovat tietoisia osallistujien kuulumisesta koe- tai kontrolliryhmään, on vaarana, että he voivat käyttäytyä eri tavalla koe- ja kontrolliryhmän osallistujien kanssa verrattaessa tilanteeseen, jossa he eivät tiedä, mihin ryhmiin osallistujat kuuluvat. Tällä voi olla vaikutusta vertailuryhmien interventioiden tai hoitomenetelmien toteutukseen, jolloin tutkimustulokset voivat olla vääristyneitä. Niiden henkilöiden sokkouttaminen, jotka toteuttavat interventiota tai hoitomenetelmää, minimoi kyseisen riskin. Sokkouttaminen tässä tilanteessa viittaa sellaisiin menettelytapoihin, jotka estävät intervention tai hoitomenetelmän toteuttajia tietämästä kummalle ryhmälle, koe- vai kontrolliryhmälle, he ovat toteuttamassa interventiota tai hoitomenetelmää.

Arvioi: Tarkista artikkelissa raportoidut tiedot niiden henkilöiden sokkouttamisesta, jotka toteuttavat interventiota tai hoitomenetelmää. Kerrotaanko artikkelissa intervention tai hoitomenetelmän toteuttajista? Säilyikö vertailuryhmien osallistujien ryhmäjako toteuttajilta salassa?

6. Ovatko tulospuuttavien mittajien sokkoutettu tutkittavien ryhmäjäoista?

Jos tulospuuttavien mittajien mittajien ovat tietoisia osallistujien kuulumisesta koe- tai kontrolliryhmään, on vaarana, että he voivat käyttäytyä eri tavalla koe- ja kontrolliryhmän osallistujien kanssa verrattaessa tilanteeseen, jossa he eivät tiedä, mihin ryhmiin osallistujat kuuluvat. Tällöin on olemassa riski, että tulospuuttavien mittaaminen vääristyy, jolloin myös tutkimustulokset voivat

olla vääristyneitä. Niiden henkilöiden sokkouttaminen, jotka mittaavat tulosmuuttujia, minimoi kyseisen riskin.

Arvioi: Tarkista artikkelissa raportoidut tiedot tulosmuuttujia mittaavien henkilöiden sokkouttamisesta. Säilyikö vertailuryhmien osallistujien ryhmäjako tulosmuuttujien mittaajilta salassa?

7. Kohdeltiinko ryhmiä yhdenmukaisesti lukuun ottamatta tutkimuksen kohteena olevaa interventiota?

Osoittaakseen 'seurauksen' olevan 'syyn' (tutkimuksen kohteena oleva interventio tai hoitomenetelmä) seurausta, olettaen ettei valintaharhaa esiinny, vertailuryhmien välillä ei tule olla mitään muita eroja toteutetun, kontrolloidun intervention tai hoitomenetelmän lisäksi. Jos 'syyn' (tutkimuksen kohteena oleva interventio tai hoitomenetelmä) kanssa samanaikaisesti ilmenee muuta altistumista, 'seurauksen' ei välttämättä voida osoittaa olevan yhteydessä tutkittuun 'syyn'. On mahdollista, että 'seuraus' voidaan tällöin selittää muulla altistumisella tai hoitomenetelmällä, jotka esiintyivät samanaikaisesti 'syyn' kanssa.

Arvioi: Tarkista raportoidut tiedot vertailuryhmien altistumisesta tai interventioista. Esiintyykö samanaikaisesti 'syyn' kanssa muita altistumisia tai hoitomenetelmiä? Onko mahdollista, että 'seurausta' voidaan selittää muilla samanaikaisesti toteutuvilla altistumisilla tai hoitomenetelmillä? Onko selvää, ettei ryhmien välillä ollut eroja muuten kuin tutkimuksessa käytetyn intervention tai hoitomenetelmän osalta?

8. Pysyivätkö tutkittavat mukana tutkimuksessa seurannan aikana, ja elleivät pysyneet, kuvattiinko ja analysoitiinko seurannan aikana ilmenneet ryhmien väliset erot asianmukaisesti?

Seurantajakso viittaa tutkimuksen ajanjaksoon, joka alkaa osallistujien satunnaistamisesta vertailuryhmiin ja jatkuu tutkimuksen loppuun saakka. Tässä arviointikriteerissä kysytään, onko kaikista tutkittavista saatavissa kaikkia tietoja (mittauksista, havainnoista, jne.) koko tutkimuksen ajalta alkaen tutkittavien satunnaistamisesta vertailuryhmiin aina tutkimuksen loppuun saakka. Metodologisessa kirjallisuudessa satunnaistamisen jälkeinen kato (eng. *post-assignment attrition*) tunnetaan ilmiönä, jossa tutkimuksessa ei ole käytettävissä tietoa kaikista mittauksista kaikkien satunnaistettujen tutkittavien osalta. RCT-tutkimukset eivät ole täydellisiä ja katoa ilmenee lähes aina. Tämän kysymyksen fokuksena on asianmukainen katoon liittyvä pohdinta (seurantajakson aikana ilmenneen kadon kuvaaminen ja mahdollisten syiden esittäminen, kadon vaikutuksen arviointi suhteessa tuloksiin jne.).

Vertailuryhmien väliset erot seurantajakson aikana esiintyvässä kadossa saattavat vaikuttaa syy-seuraus-suhteita tutkivan RCT-tutkimuksen sisäiseen validiteettiin. Erot voivat tarjota vaihtoehtoisen selityksen havaitulle 'seuraukselle' silloinkin, kun 'syy' (tutkimuksen kohteena oleva interventio tai hoitomenetelmä) puuttuu. RCT-tutkimuksen kriittisessä arvioinnissa tulee tarkistaa, ilmeneekö vertailuryhmien välillä eroja seurantajakson aikaisessa kadossa. Jos seurantajakso on puutteellinen (puutteellista tietoa tutkittavista), tarkastele yksityiskohtia raportoiduista keinoista, joita käytettiin puutteellisten tietojen käsittelemiseksi. Näitä ovat

esimerkiksi kadon kuvaaminen (tutkimuksen keskeyttäneiden määrä ja osuus, seurantajakson liittyvän kadon syyt) ja vaikutusten analyysi (analyysi kadon vaikutuksista tuloksiin).

On hyvä huomioda, että ei riitä, että tutkimuksessa kuvataan pelkästään tutkimuksen keskeyttäneiden määrää tai osuutta. Katoon liittyvien syiden pohdinta on välttämätöntä harhan riskin analysoinnissa silloinkin, kun niiden osallistujien, joista on olemassa puutteellista tietoa, määrät tai osuudet ovat samanlaiset vertailuryhmien välillä. Jos seurantajakson katoon liittyvät syyt vertailuryhmien välillä noudattavat tietynlaisia trendejä (esimerkiksi arvioitava interventio aiheuttaa sivuvaikutuksia, menetetty yhteydenpito jne.), nämä erot voivat lisätä harhan riskin todennäköisyyttä, ellei niitä tutkita asianmukaisella tavalla ja käsitellä analyysissä.

Arvioi: Onko kato kuvattu (osallistujien määrä ja syyt seurantajakson kadon selittämiseksi)? Jos vertailuryhmien välillä ilmenee eroja seurantajaksoon liittyvässä kadossa (määrä, osuus ja syyt), selvitetiinkö analyysissä katoon liittyviä trendejä? Entä, toteutetiinko analyysi, jossa tutkitaan kadon vaikutusta tutkimustuloksiin? [Huom. Kysymys 8 Ei liity hoitoaieanalyysiin (eng. Intention-to-treat) vaan tätä käsitellään kysymyksen 9 yhteydessä.]

9. Tehtiinkö lähtöryhmien mukainen (hoitoaieanalyysi eli 'intention-to-treat') analyysi?

On olemassa erilaisia tilastollisia analyysistrategioita satunnaistettujen kontrolloitujen tutkimusten aineiston analysoimiseksi, kuten hoitoaieanalyysi (tunnetaan myös englanninkielisinä käsitteenä 'intention-to-treat'), 'per-protocol' analyysi ja 'as-treated' analyysi. Hoitoaieanalyysissä osallistujia analysoidaan niissä ryhmissä, joihin heidät oli satunnaistettu riippumatta siitä, osallistuivatko he todellisuudessa näihin ryhmiin koko tutkimuksen keston ajan vai eivät, saivatko he suunniteltua koe- tai kontrolli-interventiota, tai noudattivatko he suunniteltua koe- tai kontrolli-interventiota. Hoitoaieanalyysissä verrataan alkuperäisten satunnaistettujen ryhmien osallistujien tuloksia.

Arvioi: Tarkista, raportoitiinko hoitoaieanalyysi ja tarkista analyysin yksityiskohdat. Analysoitiinko tutkittavat niissä ryhmissä, joihin heidät oli alun perin satunnaistettu, huolimatta siitä, osallistuivatko he näihin ryhmiin, ja huolimatta siitä, saivatko he suunniteltua interventiota?

[Huom. Hoitoaieanalyysi on tilastollinen analyysistrategia, jota suositellaan käytettäväksi tutkimusten raportoinnin parhaita käytäntöjä noudattavassa, raportoinnin tueksi tarkoitetussa CONSORT-ohjeistuksessa (eng. Consolidated Standards of Reporting Trials). CONSORT-ohjeistuksen mukaan hoitoaieanalyysia voidaan pitää merkinä siitä, että satunnaistetun kokeen tulosten analyysi on toteutettu metodologisesti korkealaatuisesti. Hoitoaieanalyysin avulla arvioidaan intervention vaikutusta, toisin sanoen, intervention vaikutusta silloin kun osallistujia ohjataan käyttämään interventiota. Hoitoaieanalyysi ei arvioi todellisia kyseiselle interventiolla altistumisen vaikutuksia.]

10. Mitattiinko muuttujat samalla tavalla kaikissa ryhmissä?

Jos syys-seuraussuhteita tutkivassa tutkimuksessa lopputulosmuuttujia ('seuraus') ei mitata kaikissa vertailuryhmissä samalla tavalla, tutkimuksen sisäinen validiteetti voi heikentyä. Näin

siksi, että muuttujien mittaamisessa tapahtuneet erot voivat sekoittua intervention ('syyn') vaikutuksen kanssa.

Arvioi: Tarkista, mitattiinko muuttujat samalla tavalla. Oliko sama mittari käytössä? Olivatko mittausten ajankohdat samanlaiset? Olivatko samat mittaamisen menettelytavat ja ohjeet käytössä?

11. Mitattiinko muuttujat luotettavasti?

Kausaalisia yhteyksiä tutkivassa tutkimuksessa tulosten mittaamiseen liittyvä epäluotettavuus voi heikentää 'syyn' ja 'seurauksen' välisen tilastollisen suhteen ja siihen liittyvien johtopäätösten pätevyyttä. Yksi mahdollinen virheitä selittävä tekijä tilastollisessa päättelyssä liittyen intervention ('syyn') aikaansaamaan vaikutukseen ja sen suuruuteen on lopputulosmuuttujan mittauksen epäluotettavuus.

Arvioi: Tarkista annetut tiedot kyseisen tutkimuksen mittaamisen luotettavuudesta (ei muissa lähteissä raportoidusta mittarin luotettavuudesta), kuten mittausta tehneiden lukumäärästä ja kouluttamisesta sekä saman mittajaan tekemien mittausten (eng. intra-rater reliability) ja eri mittajien tekemien mittausten (eng. inter-rater reliability) keskinäisestä luotettavuudesta. Tämä kysymys koskee kyseissä tutkimuksessa suoritettua mittauksen luotettavuutta eikä tutkimuksessa hyödynnettyjen mittareiden pätevyyttä (eng. validity). [Huom. Kaksi muuta merkittävää uhkaavaa tekijää, jotka heikentävät tilastollisten johtopäätösten validiutta ovat alhainen tilastollinen voima ja tilastollisten testien edellyttämien olettamusten sivuuttaminen. Näitä pohditaan tarkemmin kysymyksen 12 kohdalla.]

12. Käytettiinkö soveltuvia tilastollisia menetelmiä?

Epätarkoituksenmukaisen tilastollisen analyysin tekeminen saattaa aiheuttaa tilastolliseen päättelyyn liittyviä virheitä suhteessa intervention ('syyn') vaikutusten esiintyvyyteen ja laajuuteen. Alhainen tilastollinen voima ja tilastollisten testien edellyttämien olettamusten sivuuttaminen ovat kaksi merkittävää uhkaavaa tekijää, jotka heikentävät 'syyn' ja 'seurauksen' välisen tilastollisen suhteen ja siihen liittyvien johtopäätösten pätevyyttä.

Arvioi: Tarkista seuraavat asiat: noudatettiin tilastollisten testien edellyttämiä olettamuksia; suoritettiin asianmukainen tilastollinen voima-analyysi (eng. power analysis); olivatko efektikoot (eng. effect size) asianmukaiset; oliko riippuvien ja riippumattomien muuttujien lukumäärä ja tyyppi, vertailuryhmien lukumäärä ja ryhmien väliset ominaisuudet (riippumaton tai riippuvainen ryhmä) sekä tilastollisen analyysin tavoitteet (muuttujien välinen yhteys, ennuste, elinaika-analyysi jne.) huomioitu asianmukaisesti tilastollisten testien valinnassa?

13. Oliko koeasetelma tutkittavan aihealueen näkökulmasta asianmukainen, ja huomioitiinko mahdolliset poikkeavuudet perinteisestä RCT-asetelmasta tutkimuksen toteutuksessa ja analyysissä?

Joitakin koeasetelmia, kuten esimerkiksi vaihtovuoroinen koeasetelma (crossover RCT), tulee käyttää ainoastaan siinä tilanteessa, jossa sen käyttö on asianmukaista. Vaihtoehtoiset

asetelmat voivat lisätä harhan riskiä, jos niihin liittyviä tekijöitä ei oteta huomioon suunnittelu- ja analyysivaiheessa. Vaihtovuoroista koeasetelmaa tulisi käyttää ainoastaan tilanteessa, jossa tutkimukseen osallistujilla on krooninen, stabiili terveydentila, ja jossa tarkastellaan intervention lyhytaikaisia vaikutuksia (esim. oireiden helpottumista). Vaihtovuoroisen koeasetelman yhteydessä tulee varmistaa, että eri interventtioiden välissä on riittävän pitkä väli edellisen intervention vaikutusten poistamiseksi (eng. washout period).

Klusteroidussa RCT-tutkimuksessa tutkittavat satunnaistetaan ryhmiin klustereina (ns. ryppäinä, eng. clusters). Arvioitaessa klusteroidun RCT-tutkimuksen tuloksia yksilötasolla tulee ottaa huomioon analyysiyksikköön liittyviä kysymyksiä, sillä on mahdollista, että klustereiden sisällä olevien tutkittavien tulokset korreloivat keskenään. Tämä tulee ottaa huomioon analyysivaiheessa. Lisäksi ideaalia olisi, että tutkijat raportoisivat klustereiden sisäistä korrelaatiota kuvaavan kertoimen (eng. intra-cluster correlation coefficient).

Ns. "askeltaen avartuvan satunnaistamisen asetelma" (eng. stepped-wedge RCT) voi olla sopiva tilanteissa, joissa oletetaan intervention tuottavan enemmän hyötyä kuin haittaa, tai tilanteissa, jossa asetelman hyödyntämiseen päädytään esimerkiksi uuden intervention logististen, käytännöllisten tai taloudellisten seikkojen vuoksi.

Kaikki edellä mainittujen kokeiden aineiston analyysi tulee toteuttaa asianmukaisella tavalla ottaen huomioon intervention pituuden vaikutukset.

Arvioi: Käytettiinkö tutkimuksessa tutkittavaan aihealueeseen soveltuvaa koeasetelmaa? Huomioitiinko intervention pituus ja muut tekijät asianmukaisesti asetelmaa valittaessa ja tutkimusta suunniteltaessa? Huomioitiinko asetelmaan liittyvät tekijät asianmukaisesti aineiston analyysissa?