

Tutkimustiedon kriittinen arviointi

Satunnaistetun kontrolloidun tutkimuksen periaatteita ja sudenkuoppia

Sairauksien ehkäisy, hoidon ja kuntoutuksen tehon tutkimisen keskeinen työväline on satunnaistettu kontrolloitu tutkimus. Satunnaistetun tutkimusasetelman suuri valtti on tutkittavien erilaisesta valikoitumisesta tutkimusryhmiin aiheutuvan harhan hallitseminen, mutta myös satunnaistettuihin tutkimusasetelmiin liittyy sudenkuoppia. Tutkimuksen lukijan tehtävä on arvioida tulosten pätevyys, niiden kliininen merkittävyys sekä niiden mahdollinen sovellettavuus omiin hoitopäätöksiin.

Asianmukaisesti tehty satunnaistettu kontrolloitu tutkimus (randomized controlled trial, RCT) on terveydenhuollon interventioiden tehon (vaikuttavuus ihanneolosuhteissa) osoittamisen kultainen standardi. Tutkimus tulee julkaista niin selkeästi, että lukija kykenee artikkelin perusteella tekemään omat päätelmänsä tulosten luotettavuudesta, tilastollisesta merkitsevyydestä ja kliinisestä merkittävyydestä sekä yleistettävyydestä. Satunnaistetun tutkimuksen julkaisemiselle on olemassa oma, vuonna 2010 päivitetty CONSORT-ohjeistonsa (1). Satunnaistetun kontrolloidun tutkimuksen arvioinnissa huomioitavat kysymykset on koottu **TAULUKKoon 1**.

Hyvässä satunnaistetussa tutkimuksessa on selkeästi asetettu ja raportoitu PICO-periaatteen (potilas, interventio, vertailuinterventio, hoidon tulos) mukainen kysymyksenasettelu (**TAULUKKO 2**) (2). Tutkimuksen lukijan on syytä selvittää tämä itselleen, koska myös tulosten merkittävyyden ja sovellettavuuden tulkinta nojaa siihen. Erityisesti tulee pohdita, ovatko tutkimuksessa mitatut selitettävät

muuttujat potilaan kannalta merkityksellisiä. Hoitopäätöstä tehtäessä tulisi pystyä arvioimaan kaikkia potilaalle merkityksellisiä hyötyjä ja haittoja tai niiden riskiä.

Tulosten pätevyyteen liittyviä sudenkuoppia

Satunnaistetun tutkimuksen rajoitteita ovat sokkouttamisen puuttuminen tai epäonnistuminen sekä sellaiset puutteet satunnaistamisessa, että joko tutkija tai tutkittava on voinut vaikuttaa, mihin ryhmään tutkittava päätyy. Sokkouttamisen puuttumisen tai epäonnistumisen merkitys on sitä suurempi, mitä subjektiiisempia ovat tutkimuksen selitettävien muuttujien mittarit (3).

Havainnoivaan asetelmaan verrattuna satunnaistetun kokeellisen tutkimuksen heikkuutena on usein epätäydellinen sitoutuminen tutkittavaan interventioon. Esimerkiksi kirurgian vaikuttavuutta arvioivassa tutkimuksessa suuri osa konservatiiviseen hoitoon satunnaistetuista voi päätyä kirurgiseen hoitoon ja päinvastoin. Tällöin hoitoryhmien välinen ero kirurgisen hoidon saaneiden määrissä voi jäädä pieneksi ja analyysi johtaa harhaiseen tulokseen, jonka mukaan kirurgialla ei ole vaikuttavuutta. Niin ikään seurannasta pois pudonneiden suuri määrä, joka voi olla hoitoryhmien välillä epätasaisesti jakautunut, heikentää tulosten luotettavuutta. Luotettavuutta heikentää myös hoitoaieanalyysin eli lähtöryhmien mukaisen (intention-to-treat) analyysin puuttuminen. Viime aikoina on opittu tuntemaan myös tutkimuksen ennenaikaisen keskeyttämisen ja valikoivan selitettävien muuttujien raportoimisen aiheuttamat ongelmat tulosten luotettavuuteen (4).

TAULUKKO 1. Satunnaistetun kontrolloidun tutkimuksen arvioinnissa huomioitavat kysymykset (16).

Olivatko tutkimuksen tulokset päteviä?
Satunnaistettiin potilaat hoitovaihtoehtoihin asianmukaisesti?
Oliko potilaiden valitseminen hoitoryhmiin tehty siten, että potilaan tutkimukseen soveltuvuuden päättävä ja tutkimukseen hyväksyvä henkilö ei voinut tietää, mihin ryhmään potilas joutuu eli toteutuiko tutkimusryhmäjoon salassapito?
Olivatko ryhmät samanlaisia tutkimuksen alkaessa?
Olivatko potilaat, hoidon antajat ja hoitotulosten arvioijat sokkoja annetulle hoidolle?
Oliko hoitoon sitoutuminen riittävää kaikissa hoitoryhmissä?
Hoidettiin ryhmä samalla tavalla lukuun ottamatta tutkimuksen kohteena olevaa interventiota?
Oliko potilaiden seuranta riittävän pitkä ja täydellinen?
Analysoitiinko potilaat niissä ryhmissä, joihin heidät oli satunnaistettu (lähtöryhmien mukainen analyysi)?
Mitkä olivat tulokset?
Kuinka suuri oli hoidon vaikutus?
Kuinka tarkka oli arvio hoidon vaikutuksesta?
Onko tuloksista apua hoitopäätöksissä?
Voidaanko tuloksia soveltaa oman potilaani tai potilasryhmäni hoitoon?
Arvioitiinko kaikkia kliinisesti merkittäviä vaikutuksia (hoidon tuloksia)?
Ovatko hoidon todennäköiset hyödyt sen mahdollisia haittoja suurempia ja ovatko ne hoidosta aiheutuvien kustannusten arvoisia?

Satunnaistamisen onnistuminen. Potilaiden päätyminen hoitoryhmiin tulee olla satumanvaraista. Satunnaistaminen jakaa sekä tunnetut että tuntemattomat lopputulokseen vaikuttavat tekijät sattumanvaraisesti hoito- ja vertailuryhmien kesken. Jos tutkittavien lukumäärä on kyllin iso, jakauma on tasainen ja samankaltainen vertailtavissa ryhmissä. Karkeatakaan yleissääntöä riittävästä otoskoosta ei voi antaa, koska siihen vaikuttavat suuresti vallittu ensisijainen tulosmuuttuja, tutkimuksella haettavan kliinisesti merkittävän vaikutuksen suuruus ja tulosmuuttujan hajonta (2).

Potilaan valitseminen hoitoryhmiin tulee tehdä siten, että potilaan tutkimukseen soveltuvuuden päättävä ja tutkimukseen hyväksyvä henkilö ei voi tietää, mihin ryhmään potilas

tulee joutumaan eli potilaiden jakaminen tutkimusryhmiin on salattu (concealment of allocation eli tutkimusryhmäjoon salassapito). Satunnaistamisen tekevä henkilö ei saa olla mukana tekemässä päätöksiä siitä, kuka potilaista täyttää tutkimukseen oton kriteerit. Näin turvataan satunnaistamisen riippumattomuus potilasvalinnasta (2).

Sokkouttaminen. Jos hoito annetaan ilman sokkouttamista (hoitoa antava henkilö ja potilas ovat tietoisia siitä, mitä hoitoa potilas saa) ja vastemuuttujana käytetään potilaiden tai hoitavien henkilöiden omia arviointeja hoidon vaikuttavuudesta, voivat potilaiden ja hoitavien ammattilaisten odotukset ja asenteet vääristää lopputulosarviota. Mikäli selitettävä muuttuja sen sijaan on objektiivisesti mitattavissa, ei arvioinnin sokkouttamisella ole merkitystä. Potilaiden ja hoitavien henkilöiden sokkoutuksen puute voi tällöinkin heikentää tulosten luotettavuutta, esimerkiksi sen vuoksi, että potilaiden ja hoitavien henkilöiden antama oheishoito voi olla aktiivisempaa interventioon päätyneillä potilailla kuin verrokeilla (2). Sokkouttaminen voi epäonnistua esimerkiksi tilanteissa, joissa tutkittavan hoidon ja vertailuhoidon haittavaikutukset eroavat toisistaan niin paljon, että tutkimusryhmä voidaan päätellä niiden perusteella.

Hoitojen toteutuminen. Potilaiden hyvä sitoutuminen hoito- ja vertailuryhmiin tuottaa luotettavimman tiedon tehosta. Jos hoidot interventio- ja vertailuryhmissä eivät suurelta osin toteudu, voi ryhmien välille jäädä interventioissa niin vähäinen ero, että vaikuttavuutta ei ole odotettavissakaan. Samoin jos hoidon vaikutus on riippuvainen siitä, kuinka intensiivisesti ja pitkäkestoisesti hoitoa on annettu, tulee arvioida, onko tämä toteutunut siinä mittassa, että tutkimuksen tuottama arvio vaikuttavuudesta on pätevä.

Jotta satunnaistamisesta koitua hyöty säilytetään, tulee potilaat analysoida niissä ryhmissä, joihin heidät on satunnaistettu (lähtöryhmien mukainen analyysi). Esimerkiksi ver-rattaessa kirurgista ja konservatiivista hoitoa tulee konservatiivisen hoidon ryhmän potilaat analysoida siinä ryhmässä, vaikka heidät tutkimuksen kuluessa olisivatkin päädytty hoitamaan

kirurgisesti (2). On huomattava, että myös todellisessa kliinisessä tilanteessa osa potilaista kieltäytyy heille tarjotusta hoidosta tai jättää sen kesken. Lähtöryhmien mukaisen analyysin tulokset kuvastavat siis hoidon vaikuttavuutta myös arjessa.

Muiden hoitotoimenpiteiden epätasainen jakautuminen. Tutkimusryhmien hoito ei saa erota toisistaan muutoin kuin tutkittavan intervention osalta. Myös tutkimusryhmien seuranta tulee järjestää niin, että ryhmät eivät tässä suhteessa eroa toisistaan. Muut kuin tutkimuksen kohteena olevat, tutkimuksen aikana potilaisiin kohdistetut hoidot ovat lisä-interventioita, jotka pahimmillaan sotkevat koeasetelman. Näin käy, jos tutkimusryhmiin kohdistuu toisistaan eroavia lisäinterventioita (2).

Tutkimuksen keskeyttäminen. Tutkimus saatetaan eettisistä syistä keskeyttää ennen kuin tutkimussuunnitelmassa määritelty potilasmäärä on saatu täyteen. Näin tulee tehdä, jos tutkittavaan hoitoon havaitaan ennalta määritellyssä ulkopuolisen suorittamassa väli-analyysissa liittyvän vertailuhoitoon verrattuna merkittävästi enemmän vakavia haittoja. Ongelmallisempaa on, jos tutkittava hoito jo ulkopuolisen suorittamassa väli-analyysissa osoittautuu etukäteen määritellyllä tasolla tilastollisesti merkitsevästi ja kliinisesti merkittävästi vertailuhoitoa paremmaksi. Sekä teoreettisesti ajatellen että empiiristen havaintojen mukaan varhainen keskeyttäminen saattaa tällöin johtaa tutkitun intervention tehon yliarvioimiseen (5, 6, 7).

Raportoitujen tulosten valikoimisesta aiheutuva harha. Järjestelmällisten katsausten myötä useimmille on tullut tutuksi käsite julkaisuharha. Yksittäisen tutkimuksen kohdalla vastaava ilmiö on valikoiva selitettävien muuttujien raportointi. Tutkijat saattavat julkaista vain ne muuttujat, joiden osalta tulos on ollut tilastollisesti merkitsevä ja jättää joko julkaisematta tai vähemmälle huomiolle ne muuttujat, joissa tilastollista merkitsevyyttä ei ollut. Tällainen raportoitujen tulosten valikoiminen (selective outcome reporting) ja siitä aiheutuva harha on varsin tavallista ja saattaa johtaa interventioiden tehon yliarvioimiseen

TAULUKKO 2. PICO-periaatteen (potilas, interventio, vertailuinterventio, hoidon tulos) mukainen kysymyksenasettelu (2).

PICO	Määritelmä
P (patients)	Minkälaisia potilaita tutkimukseen otettiin?
I (intervention)	Mikä oli intervention (interventioiden) sisältö?
C (comparison intervention)	Mikä oli vertailuintervention sisältö?
O (outcomes)	Millä mittareilla vaikuttavuutta mitattiin (mitkä olivat ensi- ja toissijaiset lopputulosmuuttujat)?

(8, 9, 10, 11). Satunnaistetussa tutkimuksessa tulisi siis julkaista kaikki tutkimussuunnitelmassa mainitut vaikuttavuustulokset samanaikaisesti ja eritellä suunnitteluvaiheessa päätetyt ensisijaiset tulosmuuttujat toissijaisista tulosmuuttujista.

Tilastolliset menetelmät raportoitujen tulosten valikoitumisharhan osoittamiseksi puuttuvat. Cochrane-yhteisön menetelmäkäsikirjan mukaan mahdollisuutta tällaisten menetelmien saavuttamiseksi tulee selvittää vertaamalla tutkimusprotokollassa, jos sellainen on saatavilla, ilmoitettuja ensi- ja toissijaisia selitettäviä muuttujia julkaisussa raportoituihin muuttujiin. Jos tutkimusprotokollaa ei ole käytävissä, vähemmän luotettava keino on verrata julkaisun menetelmäosassa lueteltuja muuttujia raportoituihin muuttujiin. Ei-merkitsevien lopputulosten raportointi jättäminen on mahdollinen harhan lähde (12).

Saattaa olla tilanteita, jolloin tutkimuksen edetessä on osoittautunut aiheelliseksi poiketa tutkimusprotokollasta selitettävien muuttujien osalta. Näin voi käydä esimerkiksi silloin, kun osoittautuu tarpeelliseksi muodostaa useista muuttujista koostettuja yhdistelmämuuttujia. Yhdistelmämuuttujiin tulee kuitenkin suhtautua varoen, koska intervention teho muuttujan eri osatekijöihin saattaa olla erilainen (13). Yhdistelmämuuttujien tai muiden tutkimussuunnitelmaan kuulumattomien muuttujien käyttö tulee avata ja perustella tutkimusraportissa sekä raportoida niihin liittyvät tulokset toissijaisina.

Seurannasta putoaminen. Seurannasta pudonneiden hyväksyttävälle määrälle ei voida antaa absoluuttista lukuarvoa. Pudonneiden määrän merkitys liittyy sen suhteeseen loppu-tapahtumien määrästä. Mitä suurempi seurannasta pudonneiden ja mitatun lopputuloksen saaneiden tutkimushenkilöiden määrän suhde yleensä on, sitä suurempi on seurannasta putoamisen aiheuttaman harhan riski. Jos esimerkiksi 5 % tutkimuksen aloittaneista jää pois seurannasta sekä interventio- että vertailuryhmässä, mutta lopputapahtumien määrä interventio-ryhmässä on 20 % ja vertailuryhmässä 40 %, ei harhan vaaraa ilmeisestikään ole. Jos taas lopputapahtumien määrä interventio- ja tutkimusryhmässä on 2 % ja 4 %, saattaa jo 5 %:n pois pudonneiden määrä aiheuttaa huolta tuloksia tulkittaessa (4).

Seurannasta pudonneiden erilainen määrä interventio- ja vertailuryhmissä vähentää luottamusta tulosten oikeellisuuteen. Seurannasta putoaminen ei aiheuttaisi harhaa lopputulokseen, jos seurannasta pudonneiden henkilöiden – sekä interventio- että verrokkiryhmissä – todennäköisyys saada määrätty lopputulos olisi sama kuin niillä, jotka jatkoivat tutkimuksessa. Tätä ei kuitenkaan ole koskaan mahdollista varmuudella määrittää. Niinpä sekä suuri seurannasta pudonneiden määrä että erot eri ryhmien välillä heikentävät tulosten luotettavuutta.

Tuloksiin liittyviä sudenkuoppia

Tutkimusten tulokset esitetään tunnuslukujen piste-estimaatteina (arvioina vaikuttavuudesta) sekä niiden luottamusväleinä. Tunnusluvut voivat olla absoluuttisia (esimerkiksi verenpaine, kotiutumisaika) tai suhteellisia (esimerkiksi risk ratio eli riskisuhde ja odds ratio eli kerroinsuhde). Lisäksi tuloksille lasketaan tilastollinen merkitsevyys.

Piste-estimaatti tarkoittaa todennäköisintä kyseisen tutkimuksen perusteella arvioitua intervention (esimerkiksi lääkehoito tai osastokuntoutus) vaikuttavuutta vertailu-intervention verrattuna. Luottamusvälit ilmoitetaan yleensä 95 %:n mukaisesti laskettuina. Luottamusväli tarkoittaa sitä, että toistettaessa

tutkimus sata kertaa tuloksen piste-estimaatti asettuu 95 tapauksessa luottamusvälin sisälle. Tämä pätee kuitenkin vain sellaisessa potilaskokossa ja olosuhteissa, joissa tutkimus on tehty. Potilaiden ominaisuuksien kuten taudin vaikeusasteen muuttuessa kyseisessä tutkimuksessa mitattu piste-estimaatti ja sen luottamusväli eivät sen sijaan enää päde (14).

Suhteellisen riskin vähenemisellä (RRR, relative risk reduction) tarkoitetaan interventio- ja vertailuryhmän välistä eroa mitatun lopputuloksen saaneiden tutkittavien osuudessa. Jos 20 % interventio-ryhmän ja 40 % vertailuryhmän potilaista saa lopputuloksen, voidaan riski saada lopputulos merkitä luvuilla 0,2 ja 0,4. Suhteellinen riskin väheneminen ilmoitetaan yleensä prosenttilukuna, joka tässä tapauksessa on 100 % ($100 \times (0,4 - 0,2) / 0,2$). Vastavasti absoluuttinen riskin väheneminen (ARR, absolute risk reduction) tarkoittaa interventio- ja vertailuryhmän lopputuloksen saaneiden potilaiden prosenttisen määrän erotusta, joka tässä tapauksessa on 20 % ($40\% - 20\%$).

Suhteellisen riskin väheneminen pysyy usein samalla tasolla riippumatta lopputuloksen yleisyydestä. Absoluuttisen riskin väheneminen taas riippuu lopputuloksen yleisyydestä. Jos äskeisessä esimerkissä lopputulos olisi tapahtunut 2 %:lla interventio- ja 4 %:lla vertailuryhmän potilaista, olisi suhteellinen riskin väheneminen sama 100 % mutta absoluuttinen riskin väheneminen enää 2 %. Nyt päätelmä intervention hyödyllisyydestä saattaisikin olla erilainen lopputuloksen laadun (kuolema, kuumeen pitkittyminen) ja vertailujen hoitojen muiden ominaisuuksien (haitat, hinta, hyväksyttävyys potilaalle) mukaan.

Käytännön esimerkki suhteellisen riskin vähenemisestä on statiinien merkitys sydän- ja verisuonikuolleisuuden vähentäjänä (15). Statiinien aiheuttama suhteellinen merkittävän kardiovaskulaaritapahtuman riskin pieneminen viiden vuoden kuluessa on noin 20 % riippumatta henkilön SCORE-pistein arvioidun riskin suuruudesta. Absoluuttinen riskin pieneminen on suurimman riskin (yli 30 %:n riski) henkilöillä 2,18 % mutta pienimmän riskin (alle 5 %:n riski) henkilöillä 0,18 %.

Jos hoito on ollut tehokas vain jossakin tutkimusaineiston alaryhmässä, on hyöty vastaavalle potilasryhmälle epävarma. Todennäköisempää hyöty on silloin, kun erot alaryhmien välillä ovat suuria, ne on esitetty jo tutkimushypoteesissa ja sattuman riski on pieni (eli alaryhmäaineisto on kyllin suuri). Alaryhmäanalyysin tulos pitää kuitenkin yleensä todentaa muissa tutkimuksissa, ennen kuin siihen voi kunnolla luottaa (2). Lisäksi on huomattava, että jos tehoa ei ole ollut koko tutkimusaineistossa vaan vain jossain sen alaryhmässä, on hoito vastaavasti saattanut olla haitallista josain toisessa tutkimusaineiston alaryhmässä.

Tutkimuksen otoskoko tulee laskea tutkimusta suunniteltaessa. Otoskoko laskettaessa tutkija valitsee haluamansa virhetasot eron osoittamiselle (alfavirhe) tai sen jäämiselle osoittamatta (beetavirhe). Lisäksi tarvitaan arvio kliinisesti merkittävän vaikutuksen suuruudesta. Lukijan tulee tuloksia tulkitessaan ottaa kantaa siihen, onko tutkimuksessa todettu vaikutus riittävä. Jos vaikutusta ei ole osoitettu, tulee lukijan pohtia, oliko otoskoko riittävä kliinisesti merkittävän vaikutuksen osoittamiseen. Jos ei ollut, ei kyseisestä tutkimuksesta voi päätellä tutkitun hoidon tehottomuutta vertailuhoitoon verrattuna.

Lopuksi

Satunnaistetut kontrolloidut tutkimukset tehdään hyvin rajalliselle potilasjoukolla tarkoin määrittelyissä ja valvotuissa olosuhteissa. Tutkijat ovat usein hyvin erikoistuneita, osaamisalueelleen omistautuneita asiantuntijoita, ja heidän ajalliset ja muut resurssinsa on var-

mistettu tutkimuksen puitteissa. Tutkimuksiin osallistuvat henkilöt ovat hoitoon motivoituneita potilaita. Heiltä on poissuljettu useita muita sairauksia, joita terveydenhuollon arjessa vastaavaa sairautta sairastavilla potilailla puolestaan saattaa olla. Satunnaistetut kontrolloidut tutkimukset antavatkin tietoa intervention tehosta parhaassa mahdollisessa tilanteessa. Poikkeuksen muodostavat niin sanotut pragmaattiset satunnaistetut tutkimukset, joissa tutkimukseen otetaan potilaita terveydenhuollon arjesta siten, että vältetään merkittäviä poissulkukriteerejä. Selvää rajaa pragmaattisen ja ihanneolosuhteissa toteutetun niin sanotun eksploratiivisen satunnaistetun tutkimuksen välille ei voida vetää.

Satunnaistetuissa kontrolloiduissa tutkimuksissa tehokkaaksi osoitettujen interventioiden vaikuttavuus terveydenhuollon arjessa onkin usein toivottua heikompi. Interventioiden haittoja ja kustannuksia tulisikin punnita todettua tai arvioitua arkivaikuttavuutta vasten. Tähän saatetaan satunnaistettujen kontrolloitujen tutkimusasetelmien sijasta tarvita havainnoivia rekistereihin perustuvia tutkimusasetelmia, joiden tulkintaa käsitellään myöhemmin tässä artikkelisarjassa.

Tulosta voitaisiin yleensä soveltaa potilaan hoitoon, jos hän olisi tullut hyväksytyksi mukaan tutkimukseen. Käytännössä hyödyllisempää on kysyä, onko jokin tärkeä syy, miksi tuloksia ei voitaisi soveltaa juuri tähän potilaaseen tai potilasryhmään. Jos tutkimuksessa hoidolla saavutettava tulos ei potilaan mielestä ole tärkeä, tulee hoitoon ryhtymistä harkita tarkkaan yhdessä potilaan kanssa (2). ■

JORMA KOMULAINEN, LT, dosentti,
Käypä hoito -suositusten päätoimittaja
Suomalainen Lääkäriseura Duodecim, Helsinki

**PIIA VUORELA, dosentti, naistentautien ja
synnytysopin erikoislääkäri,**
Käypä hoito -toimittaja, koulutusvastaava
HUS, Porvoon sairaala
Suomalainen Lääkäriseura Duodecim, Helsinki

ANTTI MALMIVAARA, LKT, dosentti, ylilääkäri
Terveys- ja sosiaalitalouden yksikkö, Terveyden ja
hyvinvoinnin laitos (THL), Helsinki

SIDONNAISUUDET

Jorma Komulainen: Asiantuntijapalkkio (EKSSHP, PKSHP, THL, TUOTA-tutkimushanke), luontopalkkio (Lääkäriliitto)

Piia Vuorela: Ei sidonnaisuuksia

Antti Malmivaara: Ei sidonnaisuuksia

KIRJALLISUUTTA

1. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c332.
2. Straus SE, Glasziou P, Richardson WS, Haynes RB. Evidence-based medicine: how to practice and teach it. 4. painos. Edinburgh: Elsevier 2011.
3. Wood L, Egger M, Gluud LL, ym. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008; 336:601–5.
4. Guyatt GH, Oxman AD, Vist G, ym. GRADE guidelines: 4. rating the quality of evidence: study limitations (risk of bias). *J Clin Epidemiol* 2011;64:407–15.
5. Pocock SJ. When (not) to stop a clinical trial for benefit. *JAMA* 2005;294: 2228–30.
6. Montori VM, Devereaux PJ, Adhikari NK, ym. Randomized trials stopped early for benefit: a systematic review. *JAMA* 2005;294:2203–9.
7. Bassler D, Briel M, Montori VM, ym. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA* 2010;303:1180–7.
8. Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009;302:977–84.
9. Furukawa TA, Watanabe N, Omori IM, Montori VM, Guyatt GH. Association between unreported outcomes and effect size estimates in Cochrane meta-analyses. *JAMA* 2007;297:468–70.
10. Chan AW, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;291:2457–65.
11. Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ* 2005;330: 753.
12. Higgins JP, Altman D, Sterne J. Chapter 8: Assessing risk of bias in included studies. Julkaisussa: Higgins J, Green S, toim. *Cochrane handbook for systematic reviews of interventions* [verkkodokumentti]. The Cochrane Collaboration 2011. <http://handbook.cochrane.org>.
13. Ferreira-González I, Permyer-Miralda G, Busse JW, ym. Methodologic discussions for using and interpreting composite endpoints are limited, but still identify major concerns. *J Clin Epidemiol* 2007;60:651–7.
14. Guyatt GH, Oxman AD, Kunz R, ym. GRADE guidelines 6. Rating the quality of evidence: imprecision. *J Clin Epidemiol* 2011;64:1283–93.
15. Cholesterol Treatment Trialists' (CTT) Collaborators, Mihaylova B, Emberson J, ym. The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. *Lancet* 2012;380:581–90.
16. Jousilahti P, Jousimaa J, Komulainen J, Kukkonen-Harjula K, Kunnamo I, Malmivaara A. Tutkimustiedon kriittinen arviointi. Julkaisussa: Komulainen J, Honkanen M, Malmivaara A, Sipilä R, toim. *Hoitosuositusryhmien käsikirja* [verkkodokumentti]. Suomalainen Lääkäriseura Duodecim 2012. <http://www.terveysportti.fi/dtk/khk/koti>.

Summary

Principles and pitfalls of randomized controlled study

Randomized controlled study is a central tool in the study of the prevention and treatment of diseases and effect of rehabilitation. Whereas prevention of bias resulting from potential differences in study groups is a great advantage of the randomized study design, pitfalls are also involved in these studies. The task of the reader of the study is to evaluate the validity of the results, their clinical significance and possible applicability to one's own treatment decisions.

Sarjan seuraava artikkeli julkaistaan numerossa 15/2014 ja sen aiheena on:
Miten arvioin ja hyödynnän havainnoivan tutkimuksen tuottamaa tietoa?