

## Vaikuttavuustutkimus:

# Meta-analyysista lisää voimaa tulosten tulkinnan tueksi

Artikkeli on jatkoa Kasvun tuki -aikakauslehdessä julkaistulle Vaikuttavuustutkimus-artikkelisarjalle:

- Vaikuttavuustutkimus: Satunnaistettu vertailukoe tulee suunnitella huolella (2/2021)
- Vaikuttavuustutkimus: Satunnaistetun vertailukokeen laadukkaan raportoinnin edellytykset (1/2022)
- Vaikuttavuustutkimus: Miten arvioida harhaa satunnaistetuissa vertailukokeissa? (2/2022)
- Vaikuttavuustutkimus: Mitä tulee huomioida analysoitaessa satunnaistetun vertailukokeen aineistoa? (1/2023)

- Meta-analyysi on tilastollinen menetelmä, jonka avulla yhdistetään kvantitatiivisesti systemaattisen katsauksen tutkimukset.
- Meta-analyysi yhdistää useiden tutkimusten tulokset parhaaksi arvioksi näiden tutkimusten tuloksista.
- Satunnaistettujen vertailukokeiden meta-analyyseissa hyvä tutkimuskysymys määrittelee osallistujat, interventiot, vertailut ja vasteet (Participants, Interventions, Comparisons, Outcomes eli PICO).
- Meta-analyysissa on monta vaihetta:
  - 1) tutkimuskysymysten määrittäminen
  - 2) kirjallisuustietokantojen läpikäynti
  - 3) tutkimusten valinta ja tietojenhaku
  - 4) efektikokojen yhdistäminen
  - 5) heterogeenisyyden tarkastelu
  - 6) meta-analyysin tulosten raportointi.
- Yksilökohtaiseen dataan ja osajoukkoanalyysiin perustuvat meta-analyysit soveltuvat parhaiten kohdennettujen interventioiden tarkasteluun.



**SANNA  
HINKKA-YLI-SALOMÄKI**  
johtaja, biostatistiikka ja  
tiedonhallinta,  
VTL, MSc, DLSHTM  
Lastenpsykiatrian  
tutkimuskeskus,  
Turun yliopisto

### META-ANALYYSILLE ON YHTEISKUNNALLISTA TARVETTA

Tarve tehdä käytäntöön vaikuttavia päätöksiä on vauhdittanut niin sanottujen näyttöön perustuvien menetelmien käyttöönottoa. Näyttöön perustuva tutkimus voidaan määritellä järjestelmälliseksi, kvantitatiiviseksi ja ensisijaisesti kokeelliseksi lähestymistavaksi interventiosta saatavan tiedon hankkimiseen ja käyttämiseen

(Brownson ym., 2022). Tämä liittyy myös osittain tiedolla johtamisen käsitteeseen. Sen tarkoituksena on auttaa kehittämään toimintaa ja kohdentamaan voimavaroja sinne, missä niistä saadaan suurin hyöty. Tiedolla johtaminen on tiedon järjestelmällistä analysointia sen hyödyntämiseksi päätöksenteossa. Tiedolla johtamisen kokonaisuuteen kuuluvat sekä tiedon tuottaminen että tiedon hyödyntäminen. Tässä keskitymme tiedon tuottamisen yhteen osa-alueeseen, meta-analyysiin.

*”Tutkijat ovat tienneet jo vuosisatojen ajan, että yksi tutkimus ei ratkaise suurta asiaa. Itse asiassa pienen otoskoon tutkimus ei ratkaise edes pientä ongelmaa. Näin tieteen vahvana perustana pidetään tiedon keräämistä monien tutkimusten tuloksista (Hansen ym., 2022).”*

Tutkijat ovat tienneet jo vuosisatojen ajan, että yksi tutkimus ei ratkaise suurta asiaa. Itse asiassa pienen otoskoon tutkimus ei ratkaise edes pientä ongelmaa. Näin tieteen vahvana perustana pidetään tiedon keräämistä monien tutkimusten tuloksista (Hansen ym., 2022).

Meta-analyysi voidaan määritellä kvantitatiiviseksi ja muodolliseksi epidemiologiseksi koeasetelmaksi, jonka avulla voimme järjestelmällisesti arvioida aikaisempien tutkimusten tuloksia, jotta voimme tehdä päätelmiä tarkastelun kohteena olevasta tutkimusjoukosta (Haidich, 2010). Tyypillisesti meta-analyysi perustuu satunnaistettuihin vertailukokeisiin, mutta tämä ei ole välttämätöntä. Pyrkimyksenä on siis tilastollisten menetelmien avulla johtaa kokonaisarvio tai yhteisestimaatti (*pooled estimate*), joka on mahdollisimman lähellä tuntematonta totuutta. Meta-analyysia pidetään luotettavimpana todisteena näyttöön perustuvassa lääketieteellisessä ja muidenkin alojen kirjallisuudessa. Se sai alkunsa, kun haluttiin selvittää lääketieteeseen ja psykologian aloilla tehtyjen interventioiden vaikuttavuutta 1970-luvulla. Meta-analyysi on nykyisin vakiintunut menetelmä myös muun muassa johtamis- ja päätöksentekotieteissä.

*”Meta-analyyseissa on tärkeää määritellä selkeästi sisäotto- ja poissulkemiskriteerit.”*

Tavallisesti tutkijat ovat kiinnostuneita analysoimaan ja raportoimaan yhden tutkimuksen, jossa on useita vasteita. Mutta entä jos haluaisimme saada kokonaiskuvan kaikista raportoiduista tutkimuksista, joissa mukana on ollut mielenkiintomme kohteena oleva

interventio? Olisiko mahdollista yhdistää jollakin tavoin mallintamalla kaikkien tutkimusten tiedot tilastotieteellisin menetelmin? Tämän artikkelin tavoitteena on tarkastella eri tutkimuksia yhdistävän meta-analyysin ominaisuuksia ja keskustella tämäntyyppisen tutkimuksen perusteista.

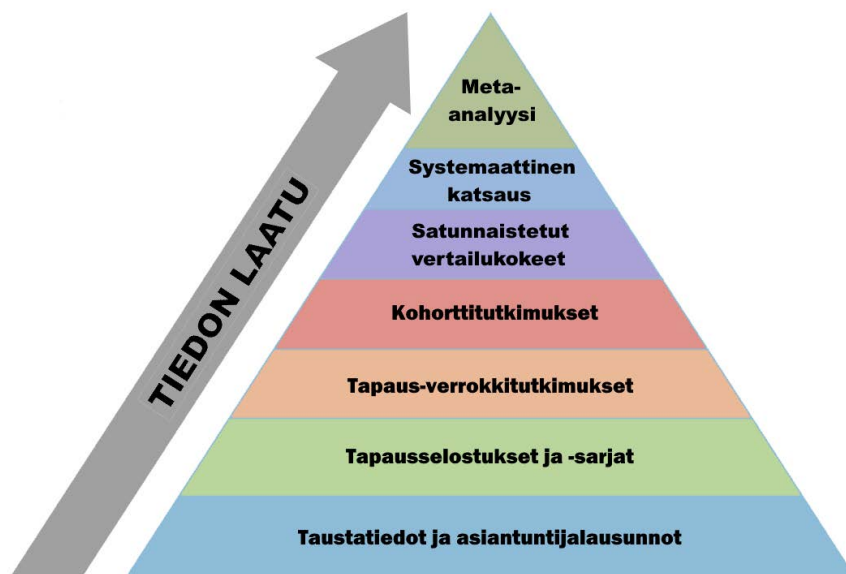
#### META-ANALYYSIT OVAT TIEDON HIERARKIAN KÄRKI

Kuvio 1 esittää tiedon laadun hierarkiaa, jossa erityyppisistä tutkimuksista saatavat tulokset on luokiteltu tutkimusta rajoittavien mahdollisten erilaisten harhojen mukaan (Cuipers, 2016). Ylimpänä hierarkiassa ovat meta-analyysit ja alimpana asiantuntijoiden lausunnot ja tapausselostukset. Tietoa kokoavat julkaisut voidaan luokitella asiantuntijoiden tekemiin perinteisiin, systemaattisiin ja meta-analyyseja sisältäviin katsauksiin. Perinteiset asiantuntijoiden kirjoittamat arvostelut painottavat kirjoittajan auktoriteettia. Yleensä on epäselvää, miten katsauksessa kuvatut tutkimukset on valittu, ja on hyvin mahdollista, että viitataan vain tutkimuksiin, jotka tukevat kirjoittajan mielipidettä.

Systemaattiset katsaukset sen sijaan noudattavat yleisiä tieteellisen tutkimuksen sääntöjä. Niillä on selkeä tavoite ja huolellisesti muotoiltu tutkimuskysymys. Tätä kysymystä tarkastellaan ennalta määritellyin kelpoisuuskeriteerein eli sen perusteella, mitkä tutkimukset valitaan katsaukseen mukaan. Tutkimuksessa noudatetaan toistettavissa olevaa metodologiaa. Katsauksissa tutkimuksia tarkastellaan kaikin osin järjestelmällisesti, arvioidaan myös tutkimusten validiteetti sekä esitetään synteesi tutkimusten tuloksista. Meta-analyysit ovat osa systemaattisia katsauksia. Systemaattisen katsauksen ominaisuuksien lisäksi meta-analyyseissä yhdistetään tulokset tilastotieteellisin menetelmin yhteen arvioon vaikuttavuudesta ja sen tilastollisesta merkitsevyydestä.

#### KUVIO 1.

##### Kerätyn tiedon hierarkia



### Miksi meta-analyseja tarvitaan?

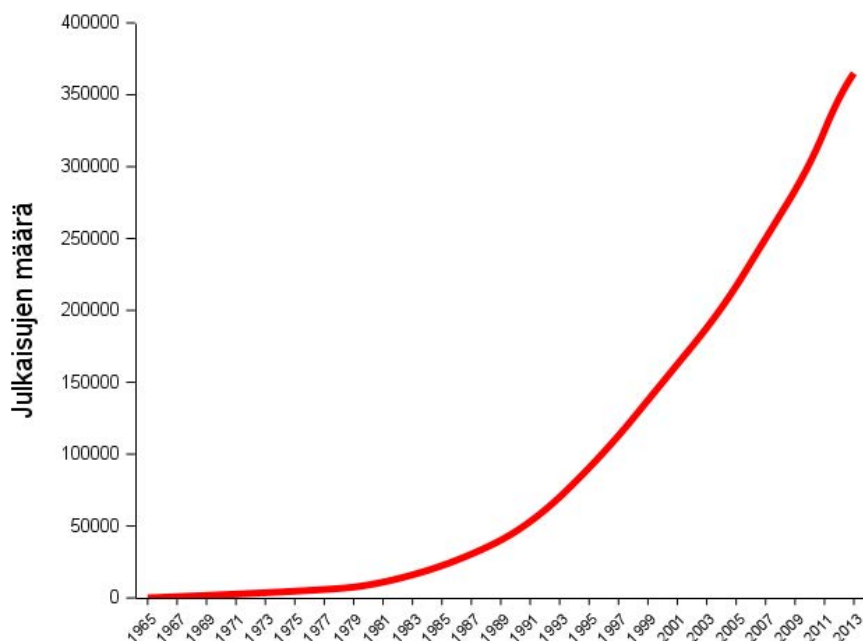
Niin kehitettyjen kuin kehitteillä olevien interventioiden vaikuttavuutta tutkitaan tyypillisesti useammin kuin kerran, usein eri tutkimusryhmissä eri paikoissa. Monissa tapauksissa näiden useiden aikapientenkin tutkimusten tulokset ovat erilaisia ja ristiriitaisia, mikä tekee yhteneväisestä päätöksenteosta vaikeaa. Myöskään mielenterveys- ja yhteiskuntatieteiden alalla satunnaistetut vertailukokeet eivät läheskään aina johda samoihin tuloksiin ja tarkastelun kohteena olevalle interventiolle lasketut efektiivisyydet (effect size) voivat vaihdella hyvinkin paljon. Tämä taas johtaa epävarmuuteen päätöksenteossa: Jos yhdessä tutkimuksessa on yhteys tutkittavien vasteiden ja

intervention välillä voimakas ja toisessa tutkimuksessa hyvin heikko, niin kumpi tulos on lähempänä todellisuutta, vai löytyykö totuus jostain näiden kahden tutkimuksen tulosten puoliväliltä?

Tieteellisissä aikakauslehdissä julkaistavien tutkimusten määrä kasvaa nopeasti. Viime aikoina on julkaistu vuosittain yhdestä kahteen miljoonaa biolääketieteellistä artikkelia (Cuipers, 2016). Suurimman biolääketieteellisen tietokannan PubMedin mukaan vuosina 1965–2013 biolääketieteen alan interventioiden vaikutuksia tutkivien satunnaistettujen vaikuttavuustutkimusten määrä kasvoi 39:stä yli 350 000:een, ja oletettavasti määrä kasvaa edelleen (Cuipers, 2016, kuvio 2).

### KUVIO 2.

Satunnaistettujen vaikuttavuustutkimusten määrä Pubmed-tietokannassa vuosina 1965–2013



Voiko meta-analyysi auttaa kohdentamaan interventiota yksilöille, jotka siitä todennäköisimmin hyötyisivät? **Thompsonin** ja **Higginsin** mukaan eniten hyötyvien ryhmien tunnistaminen ei tule koskaan olemaan helppoa, koska tarvittaisiin valtava määrä tietoa satunnaisista tutkimuksista. Jopa suurissa kokeissa, ilmisevät osajoukkojen väliset erot voivat johtua sattumasta. Isoja tutkimuksia sisältävät meta-analyysit, jotka pohjautuvat yksilökohtaiseen tutkimushenkilötietoon (*individual subject data*), mahdollistavat tutkimuskohtaisten osajoukkoanalyysien tekemisen. Tutkimuskohtaiset osajoukkoanalyysien tulokset voidaan yhdistää yli tutkimusten ja näin saadaan luotettavampia tuloksia. Niin yksittäisen henkilön hoito kuin myös valtion, kuntien tai hyvinvointialueiden päätökset tehdään ainakin jossain määrin epätäydellisen tai riittämättömän tiedon perusteella. Yksi menetelmä, joka yrittää tarkastella asiaa, on

nimeltään metaregressio (Thompson & Higgins, 2005). Menetelmässä tutkimushenkilöiden perusominaisuuksien ja taustatietojen avulla pyritään löytämään ryhmiä, jotka hyötyvät tarkasteltavasta interventiosta eniten.

*”Niin yksittäisen henkilön hoito kuin myös valtion, kuntien tai hyvinvointialueiden päätökset tehdään ainakin jossain määrin epätäydellisen tai riittämättömän tiedon perusteella.”*

## Erilaisia meta-analyyseja

Meta-analyysit voidaan luokitella kolmeen pääluokkaan (Cuipers, 2016):

- **Perinteiset tai parittaiset meta-analyysit** (*traditional, pairwise meta-analyses*):

Perinteisissä meta-analyysissä tai parittaisissa meta-analyysissä tarkastellaan yhtä vertailua, esimerkiksi eroa tarkasteltavan interventio- ja vertailuryhmän tai muun intervention välillä. Tällöin kaikki tätä vertailua tarkastelevat tutkimukset ovat mukana. Meta-analyysi arvioi sitten yhdistetyn eron tarkasteltavan interventio- ja vertailuryhmän välillä.

- **Verkkometa-analyysit** (*network meta-analyses*):

Verkkometa-analyysiin voidaan sisällyttää samanaikaisesti erilaisia vertailuja. Perinteisillä meta-analyysillä olisi suoritettava meta-analyysi jokaiselle hoidolle verrattuna vertailuryhmään. Jos useissa tutkimuksissa on tutkimushenkilöt satunnaistettu kahteen tai useampaan interventoryhmään esimerkiksi vertailuryhmän lisäksi, saadaan verkkometa-analyysissä kaikki nämä vertailut mukaan sisällyttämällä ne yhteen ja samaan analyysiin.

- **Yksilökohtaiset tutkimushenkilödatasta tehdyt meta-analyysit** (*individual subject data*):

Näissä systemaattisen katsauksen avulla identifioitujen tutkimusten yksilökohtainen data kaikista tutkimushenkilöistä on kerätty ja integroitu yhteen isoon tutkimusdataan. Monista tutkimuksista kerätty yksilökohtainen data mahdollistaa muita tarkemman arvion vaikuttavuuskoolle, koska samaa analyysimallia voidaan käyttää koko isoon aineistoon.

## Meta-analyysin etuja ja haittoja

Meta-analyysien mahdolliset edut (Cuipers, 2016):

- Meta-analyysi parantaa tutkimusvasteiden estimointitarkkuutta, sillä monet tutkimukset ovat yksinkertaisesti liian pieniä tarjotakseen vakuuttavaa näyttöä intervention vaikutuksista erikseen. Arviointi paranee yleensä, kun se perustuu suurempaan joukkoon tutkimushenkilöitä.
- Meta-analyysillä on kyky vastata tutkimuskysymyksiin, joita yksittäisissä tutkimuksissa ei pystytä tutkimaan, sekä mahdollisuus ratkaista ristiriitaisista väitteistä aiheutuvat ristiriidat. Yksittäiset tutkimukset sisältävät usein tietyntyyppisiä osallistujia ja tarkasti määriteltyjä interventioita. Valikoima tutkimuksia, joissa nämä ominaisuudet poikkeavat toisistaan, voi mahdollistaa vaikutuksen johdonmukaisuuden tutkimisen laajemmissa populaatioissa ja interventioissa. Se voi myös tarvittaessa mahdollistaa vaikutusarvioiden erojen syiden tutkimisen.
- Meta-analyysi kasvattaa tutkimuksen validiteettia, joka puolestaan mahdollistaa tutkimustulosten yleistämisen laajemmalle joukolle. Kaikissa meta-analyysin tutkimusvaiheissa pyritään mahdollisimman suureen objektiivisuuteen.
- Meta-analyysi selvittää ristiriitaisista tutkimustuloksista syntyviä kiistoja tai luo uusia hypoteeseja. Tulosten tilastollinen synteesi mahdollistaa erilaisten tulosten arvioinnin, niiden syiden tutkimisen ja kvantifioinnin.

Meta-analyysit voivat kuitenkin myös johtaa vakavasti harhaan, varsinkin jos erityisiä tutkimussuunnitelmia, tutkimuksen sisäisiä harhoja, tutkimusten välistä vaihtelua ja raportointiharhaa ei oteta huolellisesti huomioon.

Kritiikkiä meta-analyyseja kohtaan:

- **Yksi luku ei voi tiivistää mielenkiinnon kohteena olevan intervention vaikuttavuutta**

Meta-analyysi yleistää tulokset yhdeksi luvuksi, mutta jos heterogeenisyys tutkimusten välillä on merkittävää, painopisteen tulisi siirtyä yhteisvaikutuksesta itse heterogeenisyyteen. Meta-analyysi tarjoaa erilaisia työkaluja heterogeeniyyden mallin arvioimiseen ja sen mahdollisesti selittämiseen (Cuipers, 2016).

- **Verrataan eri asioita keskenään**

Meta-analyysissä yhdistetään erilaisia tutkimuksia, ja yhteisvaikutus voi jättää huomiotta näiden väliset merkittävät erot. Mikäli esimerkiksi tutkimusten sisäotto- ja poisjätökriteerit, rekrytointimenetelmä, kohdejoukko ja intervention antotapa poikkeavat tutkijoiden mielestä liikaa toisistaan, tulisi meta-analyysin tekemistä välttää (Cuipers, 2016).

- **Garbage in, garbage out**

Jos meta-analyysi sisältää heikkolaatuisia tutkimuksia, ovat tuloksetkin harhaisia ja virheellisiä. Tämän vuoksi meta-analyysin on perustuttava tarkkoihin kriteereihin sisällytettävien tutkimusten laadun suhteen (Cuipers, 2016).

- **Julkaisuharha**

Myönteisiä vaikutuksia raportoivat tutkimukset julkaistaan yleensä useammin kuin ne, jotka eivät raportoi, ja tutkimukset, joissa ei raportoida merkittäviä tuloksia, jäävät yleensä julkaisematta. Koska meta-analyysi sisältää vain julkaistuja tutkimuksia, se voi yliarvioida vaikutuksen todellisen suuruuden (Cuipers, 2016). [Julkaisuharha voidaan joskus havaita suppilokuvien \(funnel plots\) ja vastaavien tilastollisten testisuureiden, kuten Eggerin lineaarisen regressiotestin avulla.](#)

- **Kaikki vasteet eivät ole yhdistettävissä**

Toisinaan voi olla tarpeen rakentaa uusia vastemuuttujia, jotka ovat vertailukelpoisia eri tutkimusten alkuperäisiin vasteisiin (Higgins ym., 2023). Esimerkiksi yhdistettäessä tutkimuksia, joissa osassa on esitetty tulokset kategorisoidulle vasteelle ja osassa jatkuvalle vasteelle, myös jatkuvan vasteen tiedot on kategorisoitava, mikäli mahdollista, jotta vertailu on mahdollinen. Tämä vaatii aina harkintaa ja aiheen sisällöllistä asiantuntijuutta.

- **Meta-analyysi voi olla ristiriidassa satunnaistettujen vertailukokeiden tulosten kanssa**

Pääsyy meta-analyysin eroavaisuuksiin on tavallisesti se, että se perustuu heterogeenisiin ja usein pieniin tutkimuksiin. Yksittäiset tutkimukset voivat poiketa huomattavasti toisistaan esimerkiksi diagnostisten kriteerien, samanaikaisten (*comorbid*) sairauksien, taudin vakavuuden ja maantieteellisen alueen suhteen (Higgins ym., 2023).

- **Meta-analyysi käsittelee vain päävaikutuksia**  
Vaikka meta-analyysi käsittelee vain päävaikutuksia ja sen tulokset voidaan yleistää kohdeväestöön, niin yhdysvaikutuksia voidaan kuitenkin tarkastella [moderaattorianalyysin](#) avulla.

## META-ANALYYSIN TEKEMINEN

Käytännössä meta-analyysi tehdään näin:

- 1) Määritetään tutkimuskysymys, johon meta-analyysillä pyritään vastaamaan.
- 2) Etsitään tutkimukset, jotka ovat vastanneet edellämääritettyyn tutkimuskysymykseen. Nämä tutkimukset voidaan ottaa mukaan meta-analyysiin.
- 3) Kerätään tiedot sisällytettävistä ja poissuljettavista tutkimuksista [erityisesti tätä tarkoitusta varten kehitettyyn vuokaaviioon](#) ja arvioidaan tutkimusten harhan riskiä esimerkiksi käyttämällä Cochranen harhan riski -työkalua satunnaistettujen vertailukokeiden harhan arvioimisessa (Hinkka-Yli-Salomäki, 2022; Sterne ym., 2019).
- 4) Yhdistetään eri tutkimusten efektitkoot. Jos efektitkokoja ei tutkimuksissa ole ilmoitettu, ne voidaan laskea, jos tutkimuksessa on ilmoitettu muutoin riittävästi tietoja.
- 5) Tarkastellaan tutkimusten epäyhtenäisyyttä (heterogeenisyyttä) eli muusta kuin satunnaisvaihtelusta johtuvaa vaikutusten vaihtelevuutta ja mahdollista julkaisuharhaa
- 6) Kirjoitetaan ja julkaistaan meta-analyysijulkaisu.

### 1. Tutkimuskysymysten määrittäminen

Kaikki laadukas tieteellinen tutkimus alkaa hyvin muotoillusta, tärkeästä ja asiaankuuluvasta tutkimuskysymyksestä, joita voi olla yksi tai enemmän (Berlin ym., 2022). Tutkittavasta interventioista tehtyjen satunnaistettujen vertailukokeiden tuloksia yhdistävien meta-analyysien johdanto-osion tulee aina päättyä tutkimuskysymykseen, jossa määritellään osallistujat, interventiot, vertailut ja vasteet. Yhdessä näistä neljästä tutkimuskysymyksen osasta käytetään lyhennettä PICO (Participants, Interventions, Comparisons, Outcomes). Hyvä meta-analyysin tutkimuskysymys sisältää aina kaikki neljä PICO-elementtiä.

## ”Meta-analyyseissa on tärkeää määritellä selkeästi sisäänotto- ja poissulkemiskriteerit.”

Meta-analyyseissa on tärkeää määritellä selkeästi sisäänotto- ja poissulkemiskriteerit. Yksinkertaiset PICO-lyhenteen mukaiset tutkimuskysymykset eivät ole koskaan tarpeeksi tarkkoja, jotta voitaisiin päättää, mitkä tutkimukset sisällytetään ja mitkä jätetään pois. Jotta meta-analyysi olisi mahdollisimman laadukas sisällytettävien ja poissuljettavien tutkimusten suhteen, on hyvä pohtia kriteereitä mahdollisimman yksityiskohtaisesti ennen meta-analyysin aloittamista.

Psykologisten ja psykososiaalisten interventioiden tutkimuksessa myös oikean vertailuryhmän valinta satunnaistetuissa vertailukokeissa sekä meta-analyyseissä voi olla monimutkaista (Mohr ym., 2009). Tiedämme esimerkiksi, että odotuslista-tyyppiset vertailuryhmät (waiting list control groups) estimoivat todellisuutta suu-

remman intervention vaikuttavuuden verrattuna tavanomaiseen hoitoon (care-as-usual). Satunnaistettujen vertailukokeiden lisäksi meta-analyysiin voidaan sisällyttää myös ei-satunnaistettuja vertailukokeita sekä avoimia tutkimuksia. Selvää kuitenkin on, että tällaiset tutkimukset sisältävät enemmän mahdollisia harhan lähteitä eivätkä ole todistusarvoltaan yhtä luotettavia.

### 2. Kirjallisuustietokantojen läpikäynti

Psykososiaalisen ja yhteiskunnallisen tutkimuksen alalla on seuraavat kolme kirjallisuustietokantaa, joista Cuipersin (2016) mukaan tulisi ainakin etsiä:

- [PubMed](#). Tietokanta tarjoaa ilmaisen pääsyn MEDLINE-tietokantaan, joka on National Library of Medicines -tietokanta lääketieteeseen, hoitotyön, hammaslääketieteeseen, eläinlääketieteeseen, terveydenhuoltojärjestelmien ja prekliinisten tieteiden aloilta.
- [PsycInfo](#). Tietokanta on American Psychological Association -järjestön ylläpitämä maksullinen tietokanta psykologian, käyttäytymis- ja sosiaalitieteiden aloilta.
- [Cochrane Central Register of Controlled Trials, CENTRAL](#). Cochrane Collaboration ylläpitää tietokantaa, joka on useimmissa maissa maksullinen. Sisältää satunnaistetuista tai lähes satunnaistetuista vertailukokeista tehtyjä raportteja.

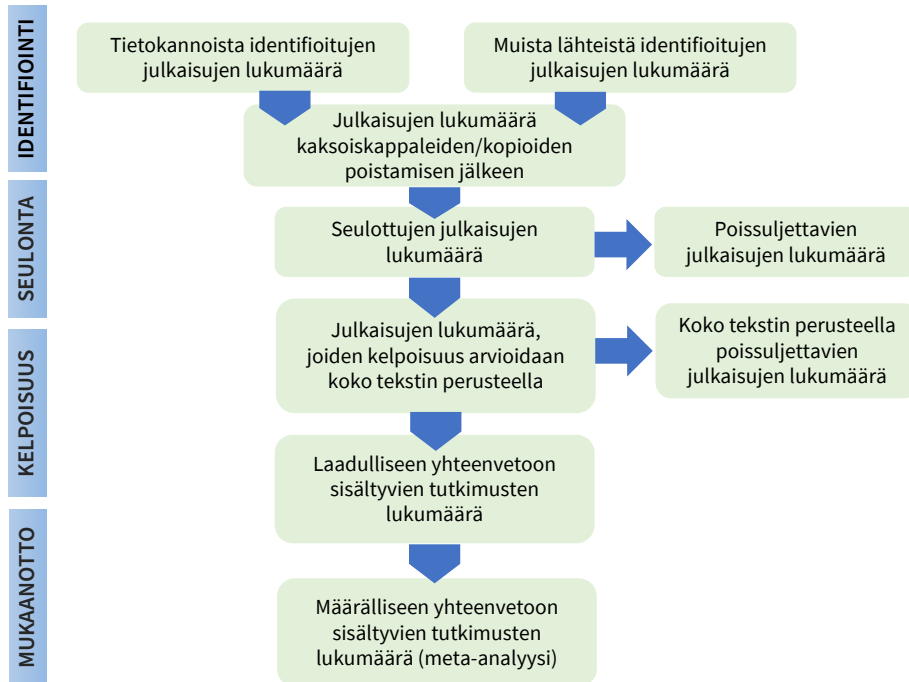
### 3. Tutkimusten valinta ja tietojenhaku

Meta-analyysiin mukaan otettavien tutkimusten valintaprosessi dokumentoidaan käyttäen PRISMA-vuokaaviota (Moher ym., 2009, kuvio 3). Vuokaavio on mukana jokaisessa laadukkaassa meta-analyysissä.

Tietojenhakuvaiheessa suosituksena on, että tutkimustyötä tekemässä olisi kaksi toisistaan riippumatonta tutkijaa. Todennäköisyys, että jokin tutkimus jää huomaamatta, on suurempi, jos vain yksi tutkija on arvioimassa mukaan otettavien tutkimusten valintaa. Mikäli tutkijoilla on toisistaan poikkeavat tutkimusvalinnat, niin ristiriidat saadaan tavallisesti ratkaistua käymällä tarkasti yhdessä läpi kyseiset julkaisut valintakriteerien suhteen.

## ”Tietojenhakuvaiheessa suosituksena on, että tutkimustyötä tekemässä olisi kaksi toisistaan riippumatonta tutkijaa.”

Valituista julkaisuista kerätään ainakin PICON mukaiset asiat eli tiedot tutkimuksiin osallistuneista, käytetyistä interventioista ja tehdyistä vertailuista. Tutkimustuloksia käytetään vaikutuskoon laskemisessa. Jokainen tutkimus sisältää harhaa, jota tutkija koettaa hallita, havaita, ennakoita ja määrittää, jotta tutkimustulokset olisivat mahdollisimman lähellä todellisuutta ja johtopäätökset voitaisiin yleistää mahdollisimman laajaan kohdeväestöön (Hinkka-Yli-Salomäki, 2022). Meta-analyysiin valittujen tutkimusten harhaa voidaan arvioida Cochranen harhan riski -työkalulla (Hinkka-Yli-Salomäki, 2022; Sterne ym., 2019). Yleisimmät harhan lähteet voidaan luokitella viiteen luokkaan: 1) valinta-, 2) toiminta-, 3) hiertymis-, 4) havaitsemis- ja 5) raportointiharha (ks. tarkemmin Hinkka-Yli-Salomäki, 2022: Miten arvioida harhaa satunnaistetuissa vertailukokeissa?). Harhojen riskin ilmoittaminen jokaiselle meta-analyysiin valitulle tutkimukselle on tärkeää. Oleellista myös on arvioida harhan osuutta yhteisissä, aggregoiduissa tuloksissa, kun kaikki meta-analyysiin valitut tutkimukset ovat analysoitu yhdessä.



#### 4. Efektikokojen yhdistäminen

Meta-analyysien tarkoituksena on yksittäisten tutkimusten tulosten yhdistäminen tilastollisten menetelmien avulla. Ennen kuin tämä on mahdollista, on eri tutkimusten tulokset harmonisoitava. Toisin sanoen, meta-analyysiin valittujen tutkimusten vasteet pyritään laskemaan jollakin valitulla standarditavalla, johon tutkimukset vain antavat mahdollisuuden. Esimerkiksi kun tehdään meta-analyysia interventioiden käytöstä masennuksen hoitoon, voi yksi tutkimus käyttää vasteena *Edinburgh Depression Scale*, toinen taas *Beck Depression Inventory* ja kolmas *Hamilton Rating Scale for Depression* -kyselyä. Näiden tutkimusten tuloksia ei suoraan voida mitenkään yhdistää, koska instrumentti on erilainen jokaisessa kolmessa tutkimuksessa.

*”Meta-analyysien tarkoituksena on yksittäisten tutkimusten tulosten yhdistäminen tilastollisten menetelmien avulla.”*

Psykososiaalisissa tutkimuksissa käytetyimmät standardoidut vaikutuskoot ovat *Cohenin d* ja *Hedgesin g* (Cuipers, 2016). Nämä vaikutusten koot perustuvat jatkuvien vasteiden tuloksiin ja osoittavat kahden ryhmän välisen eron keskihajontojen avulla. Kun psykososiaalisessa tutkimuksessa käytetään termiä efektikoko, tarkoitetaan yleensä joko *Cohenin d*:tä tai *Hedgesin g*:tä. Muita mahdollisia vaikutuskokoja ovat esimerkiksi kaksiluokkaisiin vasteisiin perustuvat suhteellinen riski (*Relative Risk, RR*) ja kerroinsuhde (*Odds Ratio, OR*).<sup>1</sup>

Efektikoon avulla voidaan tarkastella vaikutuksen suuruutta sen sijaan, että tarkasteltaisiin vain kysymystä kahden ryhmän välisestä erosta. Toisin sanoen tilastollisella merkitsevyydestillä ei arvioida vasteen vaikutuksen suuruutta vaan testataan kahden ryhmän vasteen välisen eron merkitsevyyttä. Tilastollinen testi riippuu aina ryhmien koosta, vaikutuksen koosta ja hajonnasta, jolloin suurilla otoksilla on todennäköisempää saada tilastollisesti merkitseviä tuloksia verrattuna pienempiin otoksiin. Jos *Cohenin d* ja *Hedgesin g* -arvot ovat korkeintaan 0,2, arvioidaan sen tavallisesti tarkoittavan pientä vaikutusta. Vastaavasti arvojen, jotka ovat vähintään 0,8, tulkitaan tarkoittavan suurta vaikutusta. Edellä mainittujen lukujen väliin jäävien vaikutuskokojen käsitetään tarkoittavan kohtuullista vaikutusta.

<sup>1</sup> *Cohenin d* lasketaan jakamalla ryhmäkohtaisten – tutkittavan interventoryhmän ja vertailuryhmän – keskiarvojen erotus yhdistetyllä keskihajonnalla (*pooled standard deviation*). *Hedgesin g* lasketaan samalla tavalla, paitsi keskihajonta, jossa käytetään erityisesti pienille otoksille tarkempaa otosvarianssin laskukaavaa. Tämän vuoksi on parempi aina käyttää, jos mahdollista, *Hedgesin g*:tä. Toisin sanoen *Hedgesin g* lasketaan kaavalla  $g = (m \text{ interventio} - m \text{ kontrolli}) / (\text{yhdistetty keskihajonta})$ , jossa  $m$  interventio ja  $m$  kontrolli ovat tutkittavan interventio- ja vertailuryhmän vasteen keskiarvot. Vaikutuksen koon laskemiseksi meta-analyysin artikkelit on luettava huolella, jotta tarvittavat keskiarvot ja hajonnat saadaan poimittua. Valitettavasti kaikki tutkimukset eivät kerro näitä tuloksia selvästi.

Taulukossa 1 kuvataan efektikoon (Cohenin *d*) ja NNT-luvun (interventiota saavien henkilöiden määrä, joka tarvitaan, jotta yksi henkilö hyötyisi annetusta interventiosta) välistä yhteyttä. Jos tutkimustulosten mukainen efektikoko olisi esimerkiksi 0,3 (merkitty punaisella taulukossa 1), niin meidän täytyisi hoitaa vähintään

kuutta henkilöä, jotta yksi henkilö hyötyisi tutkittavasta interventiosta. Kraemer ja Kupfer ehdottavat vaikutuksen suuruuden ja NNT-lukujen ilmoittamista aina meta-analysien julkaisuissa (Kraemer & Kupfer, 2006).

#### TAULUKKO 1.

**Cohenin *d* -efektikoon ja NNT-luvun (interventiota saavien henkilöiden määrä, joka tarvitaan, jotta yksi henkilö hyötyisi annetusta interventiosta) välinen yhteys (Kraemer & Kupfer, 2006).**

d	NNT	d	NNT	d	NNT	d	NNT	d	NNT	d	NNT
0,01	166,67	0,26	6,85	0,51	3,55	0,76	2,44	1,01	1,91	1,26	1,59
0,02	83,33	0,27	6,58	0,52	3,50	0,77	2,42	1,02	1,89	1,27	1,59
0,03	62,50	0,28	6,41	0,53	3,42	0,78	2,39	1,03	1,87	1,28	1,58
0,04	45,45	0,29	6,17	0,54	3,36	0,79	2,36	1,04	1,86	1,29	1,57
0,05	35,71	0,30	5,95	0,55	3,31	0,80	2,34	1,05	1,85	1,30	1,56
0,06	29,41	0,31	5,75	0,56	3,25	0,81	2,30	1,06	1,83	1,31	1,55
0,07	25,00	0,32	5,56	0,57	3,18	0,82	2,28	1,07	1,82	1,32	1,54
0,08	21,74	0,33	5,43	0,58	3,14	0,83	2,26	1,08	1,81	1,33	1,53
0,09	20,00	0,34	5,26	0,59	3,09	0,84	2,23	1,09	1,79	1,34	1,52
0,10	17,86	0,35	5,10	0,60	3,05	0,85	2,21	1,10	1,77	1,35	1,52
0,11	16,13	0,36	5,00	0,61	2,99	0,86	2,19	1,11	1,76	1,36	1,51
0,12	14,71	0,37	4,85	0,62	2,96	0,87	2,16	1,12	1,75	1,37	1,50
0,13	13,51	0,38	4,72	0,63	2,91	0,88	2,15	1,13	1,74	1,38	1,49
0,14	12,82	0,39	4,59	0,64	2,86	0,89	2,13	1,14	1,72	1,39	1,48
0,15	11,90	0,40	4,50	0,65	2,82	0,90	2,10	1,15	1,71	1,40	1,47
0,16	11,11	0,41	4,39	0,66	2,78	0,91	2,08	1,16	1,70	1,41	1,47
0,17	10,42	0,42	4,27	0,67	2,75	0,92	2,07	1,17	1,69	1,42	1,46
0,18	9,80	0,43	4,20	0,68	2,70	0,93	2,04	1,18	1,68	1,43	1,45
0,19	9,43	0,44	4,10	0,69	2,67	0,94	2,02	1,19	1,67	1,44	1,45
0,20	8,93	0,45	4,00	0,70	2,63	0,95	2,01	1,20	1,66	1,45	1,44
0,21	8,47	0,46	3,91	0,71	2,60	0,96	1,99	1,21	1,64	1,46	1,43
0,22	8,06	0,47	3,85	0,72	2,56	0,97	1,97	1,22	1,63	1,47	1,42
0,23	7,69	0,48	3,76	0,73	2,54	0,98	1,95	1,23	1,62	1,48	1,42
0,24	7,46	0,49	3,68	0,74	2,50	0,99	1,94	1,24	1,61	1,49	1,41
0,25	7,14	0,50	3,62	0,75	2,48	1,00	1,91	1,25	1,60	1,50	1,40

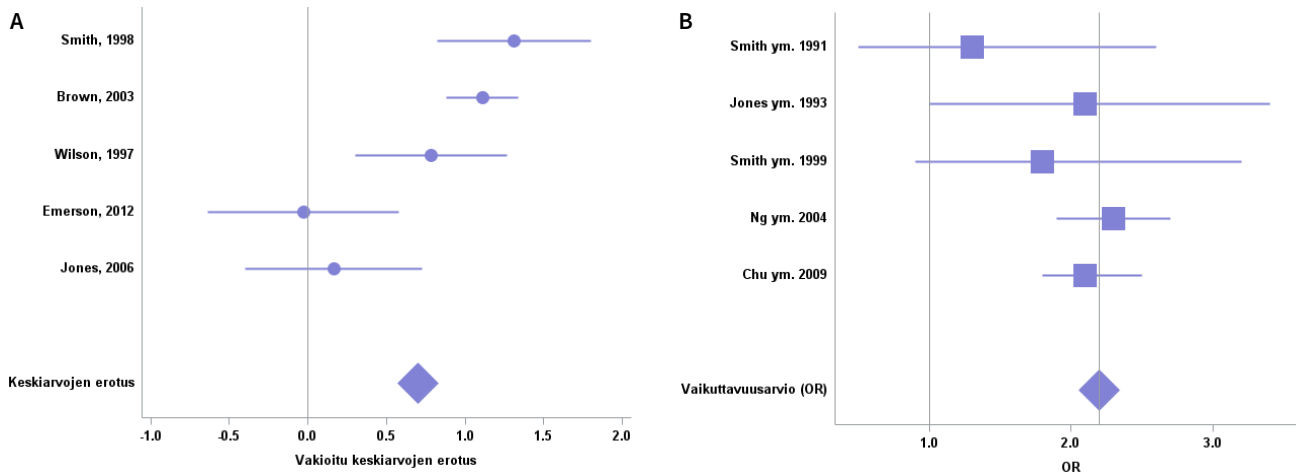
d = Cohenin *d* -efektikoko

NNT = Numbers-Needed-to-Treat, interventiota saavien henkilöiden määrä, joka tarvitaan, jotta yksi henkilö hyötyisi annetusta interventiosta

Kun jokaisesta tutkimuksesta on kerätty tarvittavat tiedot efektikokojen laskemista varten, voidaan tutkimukset yhdistää laskemalla tutkimuskohtaisten efektikokojen keskiarvo. Se kuvaa parasta mahdollista tiedossa olevaa arviota tutkittavan intervention todellisesta vaikuttavuudesta. Yksi keskeinen asia efektikokoja yhdistettäessä on tilastollisen heterogeenisyyden eli tutkimusten välisen vaihtelun huomioiminen. Tilastollisen heterogeenisyyden esiintyminen tarkoittaa, että havaitut vaikutuskoot eroavat toisistaan enemmän kuin pelkän sattuman (satunnaisen virheen) vuoksi odottaisi. Tätä voidaan estimoita ja testata tilastollisten ohjelmien avulla.

Meta-analysin tulokset esitetään tavallisesti metsikkökuvion (*forest plot*) avulla. Kuviossa 4 esitetään yksittäisten tutkimusten vaikutuskoot luottamusväleineen sekä yhdistetty vaikutuskoko luottamusväleineen (Sedgwick, 2015). Yleisesti voidaan päätellä, että mitä leveämpi luottamusvälisitä pienempi on tutkimuksen otoskoko. Kuviossa 4 esitetään esimerkit metsikkökuvioista, kun vaste on jatkuva (A) ja kaksiluokkainen (B). Seuraavaksi tarkastellaan, miten heterogeenisyyttä voidaan tutkia esimerkiksi metsikkökuvion avulla.

Esimerkki metsikkökuvioista jatkuvalle (A) ja kaksiluokkaiselle vasteelle (B)



### 5. Heterogeenisyyden tarkastelu

Heterogeenisyyttä voidaan tarkastella kolmella eri tavalla:

- 1) visuaalisesti metsikkökuvion avulla
- 2) homogeneisyydestin avulla
- 3) laskemalla indikaattoriluku  $I^2$ , joka esittää heterogeenisyyden prosenttiosuutta kokonaisvaihtelusta.

Visuaalisessa tarkastelussa saadaan selville nopeasti esimerkiksi mahdollinen yksittäinen heterogeenisyyden aiheuttama tutkimus. Esimerkiksi jos isokokaisen tutkimuksen luottamusväli ei sisälly yhteisen vaikutuskoon estimaatin luottamusväliin, voi se indikoida korkeaa heterogeenisyyttä. Kuten aiemmin mainittiin, tavallisestihan isokokaisen tutkimuksen vaikutuksen luottamusväli on kapea ja sen odotetaan sisältävän niin sanotun todellisen vaikutuskoon. Heterogeenisyyttä voidaan testata Q-testillä siten, että jos testi antaa tilastollisesti merkitsevän tuloksen, on se todiste heterogeenisyydestä (Huedo-Medina ym., 2006). Indikaattoriluvun  $I^2$  avulla voimme määrittää heterogeenisyyden suuruutta (Huedo-Medina ym., 2006). Prosenttiosuutena se vaihtelee nolasta (ei heterogeenisyyttä) sataan (kaikki vaihtelu on satunnaisvaihtelua ja interventiolla ei ole mitään osuutta). Indikaattorilukua tulkitaan siten, että 25 prosenttia indikoi matalaa, 50 kohtuullista ja 75 korkeaa heterogeenisyyttä (Higgins & Thompson, 2002).

### ”Heterogeenisyyden syiden selvittäminen on tärkeää.”

Heterogeenisyyden syiden selvittäminen on tärkeää. Voimme yrittää tehdä osajoukkoanalyseja eli poimia meta-analyysiin valituista tutkimuksista vain esimerkiksi tietyt tutkimushenkilöt jonkin taustamuuttujan suhteen. Voimme pyrkiä selvittämään tarkemmin yksittäisen poikkeavan tutkimuksen vaikutusta meta-analyysin heterogeenisyyteen sekä analysoida jatkuva vaste (kaksi- tai useampi luokkaisena) luokkamuuttujana, jos vain mahdollista.

### 6. Meta-analyysin tulosten raportoiminen

Kun kaikki edellä kuvatut vaiheet eli tutkimuskysymysten määrittäminen, kirjallisuustietokantojen läpikäynti, tutkimusten valinta ja tietojenhaku, vaikutuskokojen laskeminen ja yhdistäminen sekä heterogeenisyyden tarkastelu ovat tehty, voidaan tulokset kirjoittaa raportiksi (Cuipers, 2016). Meta-analyysia käsittelevä julkaisu noudattaa samoja sääntöjä kuin mikä tahansa muukin psykososiaalisen tutkimuksen raportti. CONSORT 2010 -ohjeiden noudattaminen ja erityisesti PRISMA-vuokaavion ja metsikkökuvion sisällyttäminen on tärkeää (Moher ym., 2010).

### COCHRANE-VERKOSTO JA -KATSAUKSET

[Cochrane-verkostolla](#) on ollut keskeinen rooli meta-analyysien laadun kehittämisessä ja tietoisuuden lisäämisessä niistä. Verkosto perustettiin vuonna 1993. Yksi sen tavoitteista on tuottaa luotettavaa ja oikea-aikaista syntetisoitua näyttöä, joka käsittelee terveyden ja hoidon päätöksenteon tärkeimpiä kysymyksiä. Cochrane-verkosto on julkaissut yli 7 500 systemaattista katsausta [Cochrane-kirjastossaan](#). Verkostolla on myös keskeinen rooli uusien menetelmien kehittäjänä tutkimustulosten syntetisoinnissa.

[Cochranen käsikirja katsauksille](#) (*Cochrane Handbook for Systematic Reviews of Interventions*) kuvaa yksityiskohtaisesti Cochrane-katsauksien valmistelu- ja ylläpitojärjestelmää terveydenhuollon interventioiden vaikutuksista. Sekä meta-analyseja että systemaattisia katsauksia suunnittelevien olisi hyvä tutustua käsikirjaan saadakseen käsityksen Cochrane-katsauksissa käytetyistä menetelmistä. Kirja sisältää ohjeita jokaiseen katsaukseen sovellettavista vakiomenetelmistä (katsauksen suunnittelu, tutkimusten etsiminen ja valinta, tiedonkeruu, harhan arviointi, tilastollinen analyysi ja tulosten tulkinta) sekä erityisemmistä aiheista (esimerkiksi ei-satunnaistetut tutkimukset, haitalliset vaikutukset, monimutkaiset interventiot, terveystaloustiede, tutkimushenkilöiden raportoimat tulokset).



## LOPUKSI

Niin yksittäisten henkilöiden hoitopäätökset kuin yhteiskunnalliset sosiaali- ja terveystieteet tehdään aina jossain määrin epätäydellisen tai riittämättömän tiedon perusteella. Jos ei esimerkiksi ole tietoa interventioiden vaikutuksista tietyissä henkilöiden osajoukoissa, niin päätökset perustuvat kaikkien tutkimuksissa mukana olleiden tutkimushenkilöiden yleiseen vaikuttavuuteen (Thompson & Higgins, 2005). Myös vertailuryhmille annetut interventiot vaikuttavat suuresti tarkastelun kohteena olevien interventioiden vaikutusten suuruuksiin. Epäsopivat vertailuryhmät voivat esimerkiksi johtaa uuden hoitomuodon yliarvioimiseen tai lupaavan kehiteillä olevan aihion lopettamiseen (Mohr ym., 2009).

Kliinisen merkittävyyden lisäksi kustannustehokkuus on huomioitava poliittisissa päätöksissä. Nykyisin monet katsovat, että hoito pitäisi olla suunnattu niille, joille se on kustannustehokainta eli henkilöille, jotka saavat suurimman kliinisen hyödyn vähimmillä resursseilla. Usein päätöksentekijöillä ei ole riittävästi tietoa päätösten perustaksi. Yksilökohtaiseen dataan perustuva meta-analyysi eli meta-analyysi, jossa tutkijalla on käytössä kaikkien tutkimusten henkilökohtainen data, tulisi Thompsonin

*”Meta-analyysin käyttö ja sen arvo todennäköisesti tulevat kasvamaan, mutta paljon riippuu tietysti myös meta-analyysin käytön onnistumisesta eli siitä, miten hyvin saamme lisättyä laadukkaan tiedon määrää.”*

ja Higginsin (2005) mukaan olla enemmän sääntö kuin poikkeus. Näin voitaisiin tarkastella paremmin muun muassa erilaisia osajoukkoja ja harvinaisempiakin häiriöitä ja oireita meta-analyysin menetelmin. Tutkijoiden mukaan tällaisten hyvin rajallisten henkilöryhmien määrittely etukäteen on välttämätöntä vääran positiivisen riskin minimoimiseksi. Yksilökohtaiseen dataan perustuvat meta-analyysit olisi tällöin suunniteltu tarkastelemaan ja löytämään parhaiten kohdennettuja interventioita.

Meta-analyysi on tehokas työkalu, kun sitä käytetään harkiten ja suunnitellusti (Higgins ym., 2023). Sen avulla saamme tiivistettyä mahdollisimman objektiivisesti olemassa olevan näytön esimerkiksi tietyntoimenpiteen vaikuttavuudesta. Meta-analyysi antaa tarkemman arvion efektiokoosta verrattuna yksittäisiin tutkimuksiin ja lisää yksittäisten tutkimusten tulosten yleistettävyyttä. Siksi sen avulla voimme saada selkeitä ja luotettavia tuloksia verrattuna yksittäisten tutkimusten antamiin keskenään ristiriitaisiin tuloksiin. Meta-analyysin soveltamisessa on kuitenkin monia vaaroja. Sen avulla tehdyt johtopäätökset ovat herkkiä mukana olevien tutkimusten metodologiselle laadulle sekä julkaisuharhalle ja kelpoisuus-kriteerien muotoilulle. Vaikka riippumattomien tutkimusten tietojen yhdistäminen meta-analyysillä voi parantaa tilastollista tarkkuutta, se ei voi kokonaan estää harhaa. Vaikka meta-analyysia kritisoidaan sen rajoituksista, ratkaisuja on olemassa: Systemaattinen lähestymistapa ja läpinäkyvyys meta-analyysin tekemisessä auttavat minimoimaan epävarmuustekijöitä sekä tekemään merkityksellisiä johtopäätöksiä. Meta-analyysin käyttö ja sen arvo todennäköisesti tulevat kasvamaan, mutta paljon riippuu tietysti myös meta-analyysin käytön onnistumisesta eli siitä, miten hyvin saamme lisättyä laadukkaan tiedon määrää. ●

*Tätä tutkimusartikkelia ei ole vertaisarvioitu.*

## AVAINSANAT:

meta-analyysi, yksilökohtainen data, yhdistetty data

## LÄHTEET:

Berlin, J. A., Rubenfeld, G. D., O’Cearbhaill, R. E., Shah, A. S. & Fihn, S. D. (2022). Keeping meta-analyses fresh. *JAMA Netw Open*. 5(8):e2228541. <https://doi.org/10.1001/jamanetworkopen.2022.28541>

Brownson, R. C., Shelton, R. C., Geng, E. H. & Glasgow, R. E. (2022). Revisiting concepts of evidence in implementation science. *Implement Sci*. 12;17(1):26. <https://doi.org/10.1186/s13012-022-01201-y>

Cuipers, P. (2016). Meta-analyses in mental health – a practical guide. Haettu 1.9.2023 osoitteesta <https://indd.adobe.com/view/5fc8f9a0-bf1e-49d3-bf5f-a40bfe5409e0>

Haidich, A. B. (2010). Meta-analysis in medical research. *Hippokratia*. 14(Suppl 1): 29–37. <https://pubmed.ncbi.nlm.nih.gov/21487488/>

Hansen, C., Steinmetz, H. & Block, J. (2022). How to conduct a meta-analysis in eight steps: a practical guide. *Manag Rev Q* 72, 1–19. <https://doi.org/10.1007/s11301-021-00247-4>

Higgins J. P. T., Thomas J., Chandler J., Cumpston M., Li T., Page M. J. & Welch V. A. (editors). (2023). *Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023)*. Cochrane, 2023. Haettu 1.9.2023 osoitteesta [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook)

Higgins, J. P. & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Stat Med*. 21:1539–1558. <https://doi.org/10.1002/sim.1186>

## LÄHTEET:

- Hinkka-Yli-Salomäki, S. (2022). Vaikuttavuustutkimus: Miten arvioida harvaa satunnaistetuissa vertailukokeissa?. *Kasvun tuki -aikakauslehti*, 2(2). Haettu 1.11.2023 osoitteesta <https://journal.fi/kasvuntuki/article/view/128906>
- Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F. & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychol Methods* 11(2):193-206. <https://doi.org/10.1037/1082-989X.11.2.193>
- Kraemer, H. C. & Kupfer, D. J. (2006). Size of treatment effects and their importance to clinical research and practice. *Biological Psychiatry* 59: 990-996. <https://doi.org/10.1016/j.biopsych.2005.09.014>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G.; PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 6(7):e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Mohr, D. C., Spring, B., Freedland, K. E., Beckner, V., Arean, P., Hollon, S. D., ... & Kaplan, R. (2009). The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychotherapy and Psychosomatics*, 78(5), 275–284. <http://doi.org/10.1159/000228248>
- Sedgwick, P. (2015). How to read a forest plot in a meta-analysis. *BMJ* 351. <https://doi.org/10.1136/bmj.h4028>
- Sterne, J. A., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., ... & Higgins, J. P. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*, 366. <https://doi.org/10.1136/bmj.l4898>
- The Cochrane Collaboration. (2020). Identifying publication bias in meta-analyses of continuous outcomes. Webinar material. Haettu 1.11.2023 osoitteesta <https://training.cochrane.org/resource/identifying-publication-bias-meta-analyses-continuous-outcomes>
- Thompson, S. G. & Higgins, J. P. T. (2005). Can meta-analysis help target interventions at individuals most likely to benefit?. *The Lancet*. 365 (9456): 341-346. [https://doi.org/10.1016/S0140-6736\(05\)17790-3](https://doi.org/10.1016/S0140-6736(05)17790-3)
- Williams, R. (2012). Moderator analyses: categorical models and meta-regression. Haettu 1.11.2023 osoitteesta [https://www.campbellcollaboration.org/images/pdf/plain-language/Moderator\\_Analysis\\_Williams.pdf](https://www.campbellcollaboration.org/images/pdf/plain-language/Moderator_Analysis_Williams.pdf)