



Ari Uusaro

Tehohoidon professori
Anestesiologian ja tehohoidon vastuukouluttaja, Itä-Suomen Yliopisto
Yliääkäri, Teho-osasto, KYS
ari.uusaro@kuh.fi

Meta-analyysin tulkinnan aakkoset

Huomioiden meta-analyysien yleistyminen lääketieteellisessä kirjallisuudessa, on tärkeää pääpiirteissään tietää, kuinka nämä analyysit tehdään ja kuinka laadukkaat meta-analyysit voi erottaa heikkolaatuisista.

Kyky ja halu omaksua uutta tietoa uran aikana on keskeinen lääkärin ominaisuus, lääketieteellinen tieto kun syntyy useimmiten uusien tutkimusten myötä. On luonnollista, kuten myös edellisen Finnanest-lehden Proffan palsta -kirjoituksessakin todettiin, ettei uusien tutkimustulosten kriittinen tarkastelu ja hyödyntäminen ole mahdollista, jos lukijalla ei ole riittävästi ymmärrystä tilastotieteestä.

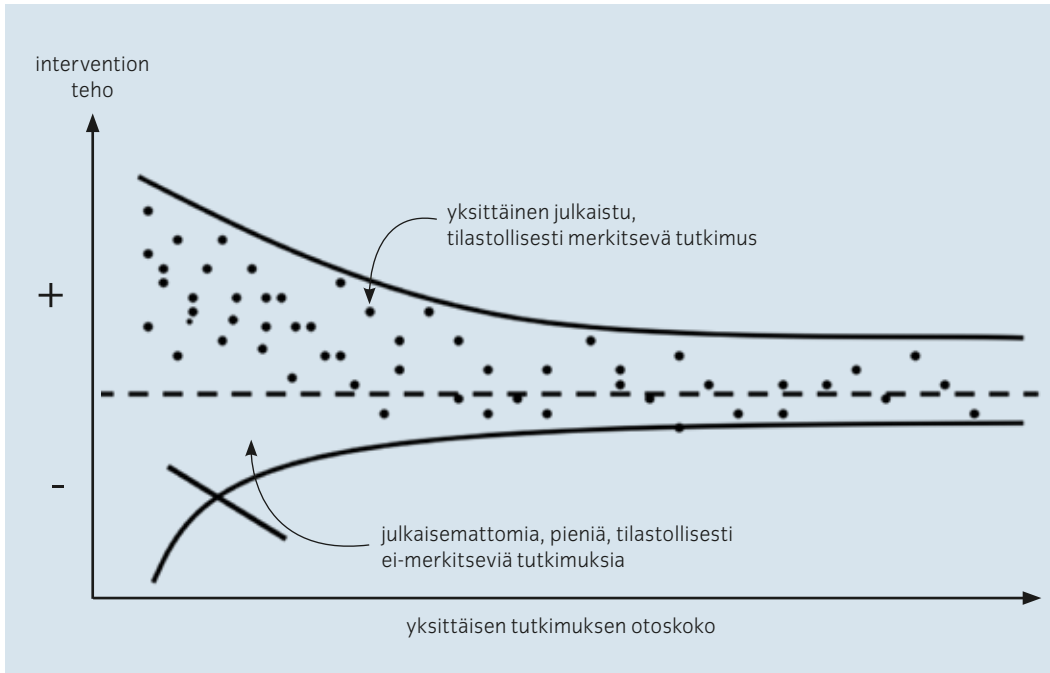
Lääketieteessä on hyvin erilaisia tutkimustyyppisiä ja mahdollisuuksia tehdä tutkimusta. On esimerkiksi kohorttitutkimuksia ilman interventiota, tapaus-verrokki (case-control) -tutkimuksia ja interventiotutkimuksia. Näiden tutkimustyyppien analysointimenetelmät poikkeavat jonkin verran toisistaan. On olemassa myös hyvin erityinen tapa tehdä lääketieteellistä tutkimusta keräämällä saman tyyppisiä, samaa asiaa selvittäviä tutkimuksia yhteen ja vetämällä näin saadusta potilasjoukosta ja käytetystä interventioista johtopäätöksiä. Kyseessä on tietysti meta-analyysi, joiden määrä on viimeisten kahden vuosikymmenen aikana meidänkin erikoisalalamme lisääntynyt valtavasti. Luonnollisesti lukijalta vaaditaan kriittisyyttä näidenkin tutkimusten lukemisessa ja tulkinnassa. Kuten yksittäisiä julkaistuja tutkimuksia, myös meta-analyysit on hyviä ja huonoja. Tämän kirjoituksen tavoitteena on auttaa lukijaa arvioimaan meta-analyysien sisältöä ja erottamaan laadukkaat meta-analyysit

vähemmän laadukkaista (1). Kaikkia meta-analyysien tulkintaan liittyviä yksityiskohtia ei luonnollisestikaan ole mahdollista yhden artikkelin puitteissa käydä läpi.

Miksi meta-analyysit tehdään?

Riittävän suurella potilasmäärällä tehty satunnaistettua kontrolloitua tutkimusta (randomized controlled trial, RCT) pidetään luotettavimpina tapana osoittaa hoitointervention mahdolliset hyödyt, haitat tai vaikuttamattomuus ideaaliympäristössä. Usein kuitenkin ei ole olemassa yhtä tai useampaa riittävän suurta yksittäistä satunnaistettua interventiotutkimusta, jotta hoidon teho pystyttäisiin riittävän luotettavasti osoittamaan tai sulkemaan pois. Joskus voi olla kyse pelkästään siitä, että yksittäisen interventiotutkimuksen otoskoko on niin pieni, ettei tutkimuksen voima riitä eri hoitoryhmien välisen eron osoittamiseen, vaikka tällainen ero todellisuudessa olisi olemassa.

Meta-analyysissä on kysymys tutkimusmenetelmästä, jossa tilastollisia keinoja käyttäen yhdistetään useita samaa aihetta selvittäviä tutkimuksia tilanteessa, jossa yksittäiset tutkimukset ovat olleen liian pieniä mahdollisen todellisen eron löytämiseksi hoitojen välillä (jos se on olemassa), tai sen osoittamiseksi, ettei eri interventioiden välillä ole kliinisesti merkitsevää eroa. Useita pienempiä tutkimuksia yhdistämällä meta-analyysissä tutkimuksen voima kasvaa, ja näin saadaan



Kuva 1. Julkaisuharha meta-analyysissä. Tutkimus julkaistaan todennäköisemmin, jos siinä on tilastollisesti merkitsevä löydös. Erityisesti pieni tutkimus julkaistaan todennäköisemmin, jos löydös on tilastollisesti merkitsevä ja positiivinen. Tämä voi aiheuttaa sen, että meta-analyysissä intervention teho yliarvioidaan, jos siinä on mukana paljon pieniä tutkimuksia.

täsmällisempi arvio (tuloksen piste-estimaatti ja tavallisesti 95 % luottamusväli) tutkittavan intervention tehosta tai tehottomuudesta.

Piste-estimaatilla tarkoitetaan tutkimuksen tulosta – täsmällistä lukua tutkimustulosta kuvaamaan. Piste-estimaatti voi olla vaikkapa relatiivinen riski jollekin tapahtumalle tai ero kuolleisuudessa prosentteina eri ryhmien välillä. Piste-estimaatin ympärillä oleva luottamusväli kuvastaa piste-estimaatin tarkkuutta tai varmuutta tutkimustuloksen kuvaamisessa. Tutkimustuloksen varmuuden parantuminen näkyy tuloksissa siten, että piste-estimaatin ympärillä oleva tuloksen luottamusväli kapenee.

Meta-analyysillä on ainakin teoriassa mahdollista löytää kliinisesti merkittävä potilashoitoa potentiaalisesti parantava uusi interventio, ennen kuin riittävän suuri yksittäinen satunnaistettu tutkimus on tehty. Uusi hyödyllinen hoito olisi näin ollen mahdollista saada aikaisemmin käyttöön. Meta-analyysijä on kuitenkin myös kritisoitu. On joskus käynyt niinkin, että samaa interventiota tutkiva meta-analyysi ja myöhemmin samasta aiheesta tehty hyvin suuri yksittäinen satunnaistettu kontrolloitu tutkimus voivat päätyä eri johtopäätöksiin intervention hyödyllisyyden suhteen. Tämänkin vuoksi on tärkeää,

että lukija tuntee pääpiirteissään meta-analyysien tekotavan ja sen, mitkä voivat olla potentiaalisia virhelähteitä yksittäisissä meta-analyysissä (1).

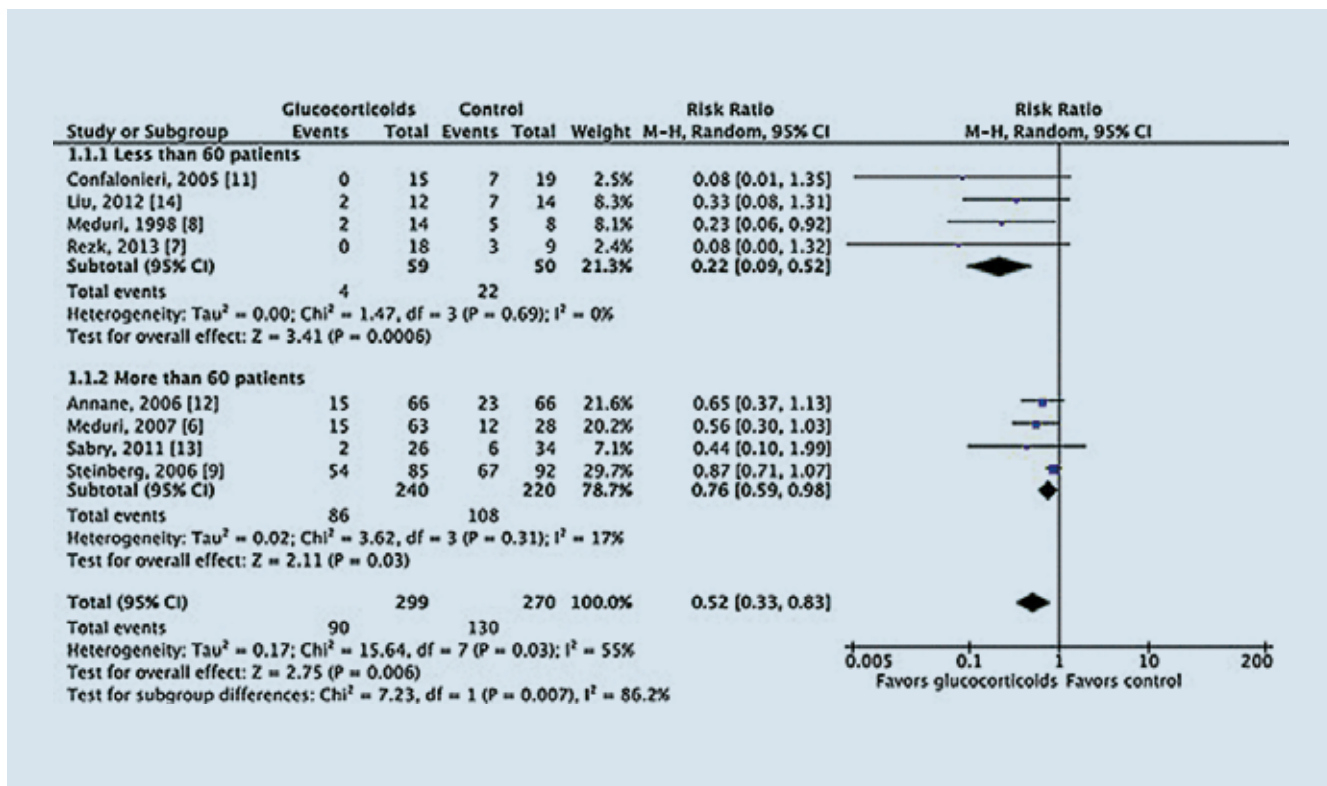
Tutkimuskysymyksen muotoilu ja relevanttien laadukkaiden tutkimusten löytäminen

Meta-analyysissä, kuten kaikessa tutkimuksessa, ensimmäinen tehtävä on muotoilla kliinisesti relevantti ja täsmällinen tutkimuskysymys. On tärkeää määrittellä täsmällisesti, mitä potilaita tutkimukseen otetaan (esimerkiksi ARDS-potilailla taudin vaikeusaste), mikä täsmällisesti on tutkittava interventio (esimerkiksi ARDS-potilailla tietty lääke tai hengityslaitteen säätö), mitä päävastemuuttujaa tutkitaan (esimerkiksi täsmällisesti määriteltä kuolleisuus tai hoitoajat). Samoin on määriteltävä, mikä on vastaava hoito verrokkiryhmässä (esimerkiksi PEEP-taso, lääketutkimuksessa lumelääke). Mitä täsmällisemmin hoitoon ja erityisesti hoitointerventioon liittyvät seikat on kuvattu, sitä selkeämpää on meta-analyysin tulosten arviointi.

On oleellista pystyä täsmällisesti kuvaamaan ne kriteerit, joiden mukaan yksittäiset tutkimukset on otettu mukaan meta-analyysiin.

Meta-analyysissä yhdistetään useita samaa aihetta selvittäviä tutkimuksia.

>>



Kuva 2. Sairaalakuolleisuus ARDS-potilailla (n = 569), jotka on satunnaistettu saamaan pitkitettyä kortikosteroidihoitoa tai rutiinihoitoa ilman kortikosteroideja alle 14 vuorokautta ARDS:n alusta. Yksittäisiä tutkimuksia on 8 ja tutkimukset on jaettu kahteen ryhmään sen perusteella, onko tutkimuksessa ollut enemmän tai vähemmän kuin 60 potilasta. (2)

Julkaisuharha voi vääristää merkittävästi meta-analyysin tuloksia.

Tutkijoiden täytyy myös olla samaa mieltä näiden yksittäisten tutkimusten sisällyttämisestä meta-analyysiin. On hyvin tärkeää kuvata täsmällisesti hakumenettelyt, joilla yksittäisiä tutkimuksia on etsitty. Mitä tietokantoja on käytetty, onko hakuprosessissa oltu yhteydessä teollisuuden (mahdolliset julkaisemattomat tutkimukset), onko oltu yhteydessä yksittäisiin tutkijoihin? Mahdolliset kielirajoitteet on syytä huomioida. Useimmat meta-analyysit on tehty vain englanninkielisistä julkaisuista, mutta voi hyvin olla, että joitakin merkittäviä tutkimuksia on julkaistu vaikkapa vain saksankielisessä kirjallisuudessa.

Julkaisuharha ja pienet tutkimukset

Julkaisuharha (= publication bias) voi vääristää merkittävästi meta-analyysin tuloksia. Julkaisuharha tarkoittaa sitä, että tilastollisesti merkitsevät tutkimukset tulevat helpommin julkaistuksi kuin tilastollisesti ei-merkitsevät tutkimukset. Jos tilastollisesti ei-merkitseviä, mutta metodo-

logisesti laadukkaita julkaisematta jääneitä tutkimuksia ei oteta mukaan meta-analyysiin, saattavat tulokset vääristyä. Tiedetään myös, että tilastollisesti merkitsevät tutkimukset tulevat julkaistuksi aiemmin kuin tilastollisesti ei-merkitsevät tutkimukset.

On syytä huomioida, että jos tutkimus on otoskooltaan pieni ja sen tulos on tilastollisesti merkitsevä, niin tilastotieteen perusteista johtuen tutkimuksessa käytetyn lopputuleman eron (esimerkiksi ero kuolleisuudessa) eri hoitoryhmien välillä täytyy pienessä tutkimuksessa olla suuri tullakseen tilastollisesti merkitseväksi. Jos meta-analyysissä on lukuisia pieniä tutkimuksia, niin ero hoitoryhmien välillä voi tehdyssä meta-analyysissä keinoitekoisesti suurentua julkaisuharhasta johtuen. Otoskooltaan pienistä julkaisemattomista negatiivisista tutkimuksista johtuvaa julkaisuharhaa meta-analyysissä pyritään selvittämään kuvassa 1. Julkaisuharhan havaitsemiseksi on olemassa myös tilastollisia menetelmiä.

Yksittäisten tutkimusten laadun arviointi ja tutkimusten heterogeenisyys

Huonojen satunnaistettujen kontrolloitujen tutkimusten yhdistäminen meta-analyysiksi tuottaa vain epäluotettavia tuloksia. Tästä syystä on tärkeää arvioida meta-analyysissä olevien yksittäisten tutkimusten metodologinen laatu. Yksittäisten tutkimusten laadun arviointiin on olemassa lukuisia menetelmiä, ja tutkimusten laatu voidaan vaikkapa pisteyttää. Tällaisten pisteytysmenetelmien käyttöön voi kuitenkin liittyä ongelmia. Onkin tärkeää, että tutkimuksen laadun kannalta oleelliset seikat on huomioitu, kun sen laatua arvioidaan: onko satunnaistamismenetelmä (= randomointi, randomisointi) ollut asianmukainen, onko tulokset analysoitu sokkoutetusti, onko tulosten analyysi tehty ns. ”intention-to-treat”-menetelmällä, onko potilaiden seuranta ollut kattavaa ja miten tutkimuksesta mahdollisesti ennenaikaisesti poistuneet potilaat on huomioitu?

Oleellisen tärkeää meta-analyysin laatua tarkasteltaessa on arvioida, onko yksittäisten tutkimusten tuloksia järkevää yhdistää. Poikkeavatko esimerkiksi tutkittavat potilaat tai tutkittava interventio eri tutkimuksissa niin paljon, että tutkimusten yhdistäminen ei ole mielekästä. Tämän asian arvioinnissa pitää käyttää erityisesti kliinistä päättelyä. Tulosten yhdistettävyyttä arvioidessa käytetään heterogeenisyys-termiä: ovatko tutkittavat potilaat ja/tai interventio mahdollisimman samanlaisia, eli tutkimusten heterogeenisyys on pieni? Jos yhdistetään hyvin heterogeenisiä tutkimuksia, niin meta-analyysin tulos ei ole luotettava. Jos meta-analyysissä tutkitaan esimerkiksi ravitsemushoidon vaikuttavuutta sairaalapotilailla, ei ole metodologisesti oikein yhdistää tutkimuksia, joissa osassa on mukana jo lähtökohtaisesti aliravittuja potilaita ja osassa potilaiden ravitsemustila on jo lähtökohtaisesti ollut normaali. Jos tutkimusten välillä on heterogeenisyyttä, on yksi meta-analyysin laadintavaihtoehto analysoida erilaiset potilasryhmät toisistaan erillään.

Mahdollisen heterogeenisyyden havaitsemiseksi on olemassa myös tilastollisia menetelmiä. Yksi usein käytetty on ns. I^2 -testi, jossa I^2 voi saada minkä tahansa arvon välillä 0 – 100 %. Tällöin 0 % tarkoittaa, että heterogeenisyyttä ei ole lainkaan, ≥ 50 % tarkoittaa, että heterogeenisyyttä on kohtalaisesti ja ≥ 75 % tarkoittaa, että heterogeenisyyttä on runsaasti. 100 % tarkoittaa,

että heterogeenisyys on mahdollisimman suurta, mikä on potentiaalisesti suuri ongelma meta-analyysin tulosten tulkinnassa.

Tulosten esittely ja pohdinta

Laadukkaassa meta-analyysissä esitellään ns. ”flowchart” (vuokaavio, prosessikaavio), jossa kuvataan, kuinka meta-analyysiin mukaan otetut yksittäiset artikkelit on löydetty ja millä perusteella artikkeleita on mahdollisesti karsittu pois. Tuloksia esitettäessä yksittäiset analyysissä mukana olevat artikkelin esitellään tavallisesti hyvin lyhyesti (esimerkiksi otoskoko, interventio, tulos). Tämän jälkeen meta-analyysin tulos esitetään yleensä graafisesti. Yksittäisten tutkimusten tuloksen piste-estimaatti ja tavallisimmin 95 %:n luottamusväli esitellään, jonka jälkeen esitetään päälöydöksen piste-estimaatti ja estimaatin luottamusväli, kun eri tutkimusten tulokset on yhdistetty. Esimerkki yksittäisten tutkimusten tulosten yhdistämisestä meta-analyysissä on esitetty kuvassa 2. Kuvan esimerkissä tutkimusten yhdistämisen jälkeen todetaan, että kortikosteroidien anto (verrattuna lumelääkkeeseen) vähentää sairaalakuolleisuutta (RR, risk ratio = 0,52, 95 %:n luottamusväli 0,33–0,83). Kun kaikki yksittäiset tutkimukset on yhdistetty, voidaan myös todeta, että I^2 -arvo on 55 %, mikä viittaa siihen, että tutkimusten välillä on kohtalaisesti heterogeenisyyttä (2).

Meta-analyysin diskussioon eli pohdintaoisioon kuuluu tutkimuksen mahdollisten heikkouksien tai epävarmuuden arviointi samaan tapaan kuin muissakin tutkimuksissa on tapana tehdä. Lopuksi arvioidaan esimerkiksi tulosten yleistettävyyttä ja mahdollista vaikutusta nykyisiin hoitokäytäntöihin tai tuleviin tutkimustarpeisiin. ■

Viitteet

1. PRISMA-statement (Preferred Reporting Items for Systematic Reviews and Meta-analysis). <http://www.prisma-statement.org/> (luettu 06.03.2020)
2. Meduri GU, Bridges L, Shih M-C, ym. Prolonged glucocorticoid treatment is associated with improved ARDS-outcomes: analysis of individual patients' data from four randomized trials and trial-level meta-analysis of the updated literature. *Intensive Care Med* 2016; 42: 829–40.

Meta-analyysiä tarkasteltaessa tulee arvioida yksittäisten tutkimusten tulosten yhdistämisen mielekkyyttä.