

Lasten ja nuorten mielenterveyttä edistävien psykososiaalisten menetelmien vaikuttavuuden arviointi

- Vaikuttavuustietoa saadaan satunnaistetuilla vertailututkimuksilla (*randomized controlled trial*, RCT) tai vertaiskontrolloiduilla tutkimuksilla (*benchmarking controlled trial*, BCT). Lasten ja nuorten mielenterveyttä edistävien psykososiaalisten menetelmien (interventioiden) arvioinnissa voidaan hyödyntää näitä molempia.
- Yksilöön kohdistuvissa yksittäisissä menetelmissä (kliininen impaktitutkimus) satunnaistettu vertailukoetutkimus on yleensä luotettavin menetelmä, ja vertaisarvioinnissa vertaiskontrolloitu tutkimus on ainoa käytettävissä oleva menetelmä.
- Järjestelmään kohdistuvien menetelmien vaikuttavuutta (järjestelmäimpaktitutkimus) arvioidaan joko ryhmäsatunnaistetuilla vertailututkimuksilla tai vertaiskontrolloiduilla tutkimuksilla.
- Pätevä arvio tulosten sovellettavuudesta arjen oloihin edellyttää kattavaa kuvausta tutkittavien ominaisuuksista, vertailussa olevista interventioista sekä tuloksista. Järjestelmäimpaktitutkimuksissa tarvitaan lisäksi kuvaus verrattavien järjestelmien ominaisuuksista.
- Vertaiskontrollonin menetelmällä (*benchmarking-method*) voidaan arvioida satunnaistettuja vertailututkimuksia ja vertaiskontrolloituja tutkimuksia sekä näistä tehtyjä järjestelmällisiä katsauksia.



ANTTI MALMIVAARA
professori, lääketieteen ja
kirurgian tohtori

Iltan Kasvun tuen tavoite on arvioida tarjotun tuen tai hoidon vaikuttavuutta tilanteissa, joissa pyritään edistämään lasten tai nuorten mielenterveyttä ja psykososiaalista hyvinvointia tai ehkäisemään ja hoitamaan varhaisessa vaiheessa joko suoraan tai välillisesti lasten ja nuorten mielenterveysongelmia. Kohderyhmänä ovat alle 18-vuotiaat lapset kasvu- ja kehitysympäristöissään, esimerkiksi koulussa, kotona tai harrastuksissa. Eettisenä arvopohjana ovat asiakaslähtöisyys, osallisuus, voimaannuttavuus ja vuorovaikutuksellisuus.

Tämän artikkelin tavoitteena on kuvata vaikuttavuuden arviointiin liittyviä peruseriaatteita sekä niiden soveltamista lasten ja nuorten mielenterveyttä edistävien psykososiaalisten menetelmien (interventioiden) vaikuttavuuden arvioinnissa.

SOSIAALI- JA TERVEYDENHUOLLON TAVOITTEET

Sosiaali- ja terveydenhuollon päätavoitteet ovat sairauksien ehkäisy ja parantaminen, elinajan pidentäminen, oireiden lievittäminen, toimintakyvyn ylläpitäminen ja kohentaminen sekä yksilöllisen ja yhteisöllisen hyvinvoinnin edistäminen (Malmivaara, 2022a). Palvelujen saatavuus, laadukas toteuttaminen sekä yhdenvertaisuus laadukkaisiin palveluihin pääsemisessä ovat vaikuttavuuden, turvallisuuden ja kustannusvaikuttavuuden edellytys (Malmivaara, 2020). Iltan Kasvun tuen tavoite on tukea ja täydentää tätä toimintaa arvioimalla systemaattisin katsauksin lasten ja nuorten mielenterveyttä edistävien psykososiaalisten menetelmien vaikuttavuutta sekä edistämällä vaikuttavien menetelmien implementointia Suomessa.

VAIKUTTAVUUS JA ARKIVAIKUTTAVUUS

Vaikuttavuuden arvioinnin lähtökohtana on selvittää, mitä lisäarvoa asiakkaalle voidaan saada jollakin menetelmällä (toimintatavalla) (Moher ym., 2010; Malmivaara, 2022a). Vaikuttavuuden arvioimiseksi tarvitaan vertailuryhmä, jossa toimii ei ensinkään ryhdytä, käytetään aiempaa toimintatapaa tai toimitaan jollain vaihtoehtoisella tavalla (Morabia, 2006). Vaikuttavuus on menetelmän aiheuttama muutos saaduissa tuloksissa, kun menetelmää käytetään tietyn määritellyn vaihtoehdon sijaan. Kyse on tuotettujen tulosten erotuksesta asetelmassa, jossa vaikuttavuuden määrittämiseen on käytetty asianmukaista epidemiologista asetelmaa ja tilastotieteen menetelmiä. Pelkkä hoitotulosten mittaaminen ennen

ja jälkeen asiakkaan saamia toimenpiteitä jättävät vaikuttavuuden tuntemattomaksi (*unknowable*). (Miettinen ym., 2019; Miettinen, Henkilökohtainen tiedonanto, 2020.)

”Vaikuttavuus on menetelmän aiheuttama muutos saaduissa tuloksissa, kun menetelmää käytetään tietyn määritellyn vaihtoehdon sijaan.”

Vaikuttavuus on yläkäsite, jonka alle sijoittuvat kokeellinen vaikuttavuus ja arkivaikuttavuus Malmivaara ym., 2022). Kokeellisesta vaikuttavuudesta saadaan tietoa satunnaistetuilla vertailukoikeilla (*randomized controlled trial*, RCT), joissa tutkittavat arvotaan kahteen tai useampaan ryhmään (Moher ym., 2010). Satunnaisen ryhmiin jakautumisen ansiosta riittävän suurilla tutkittavien asiakkaiden määrillä päästään keskimääräisesti ominaisuuksiltaan samankaltaisiin tutkimusryhmiin. Kun satunnaistettujen vertailukoetutkimusten tuloksia sovelletaan arjen olosuhteisiin, huomioidaan asiakasaineistoissa, käytetyissä menetelmissä (interventioissa) ja vastemuuttujissa (tulostittareissa) olevat samankaltaisuudet ja eroavuudet. Tämän vertailun perusteella tehdään oletukset arkivaikuttavuudesta (Malmivaara ym., 2022b).

Arkivaikuttavuudesta saadaan suoraa tietoa vertaiskontrolloiduilla tutkimuksilla (*benchmarking controlled trial*, BCT), joissa lähtökohtaisesti valitaan ominaisuuksiltaan vertailukelpoiset tutkittavat ryhmät, ja lisäksi alkutilanteessa olevat ryhmien väliset erot vakioidaan tilastollisilla menetelmillä (Malmivaara, 2015; Simula ym., 2024). Näillä toimilla pyritään mahdollisimman hyvään ryhmien vertailukelpoisuuteen alkutilanteessa. Vertaiskontrolloitujen kokeiden tavallinen tutkimuskohde on asiakkaita auttavien yksiköiden keskinäinen paremmuus saavutetuissa tuloksissa. Mikäli alkutilanteen vakioinnissa on onnistuttu, johtuu tulosten erotus siitä, kuinka paljon paremmin hyvät tulokset saanut yksikkö on toimenpiteissään onnistunut.

PICOS- JA CAIMER-VIITEKEHYKSET

Vaikuttavuustutkimuksen tai -arvioinnin kysymyksenasettelu muotoillaan siten, että käy ilmi, minkälaisesta kohderyhmästä ja minkälaisesta interventioista (menetelmästä) on kyse sekä millä tulostittareilla (vastemuuttujilla) vaikuttavuutta arvioidaan. Eriyisesti terveydenhuollon vaikuttavuuden arviointiin on kehitetty PICOS-viitekehys, jonka nimi muodostuu englanninkielisistä sanoista *Population, Intervention, Control intervention, Outcome* sekä *Study design* (Moher ym., 2010).

Vaikuttavuuden arvioinnissa tarvitaan sekä objektiivisia että subjektiivisia mittareita. Jälkimmäisiä tarvitaan, koska terveydentila perustuu merkittävältä osalta yksilön kokemukseen. Kokemusta taudin aiheuttamasta haitasta voidaan mitata niin sanotuilla PROM-mittareilla (*patient related outcome measures*). PROM-mittarit voivat olla tauti- tai tilakohtaisia (esim. selkäsairauksien Oswestry-toimintakyvyindeksi) tai yleisiä terveyteen liittyvän elämänlaadun mittareita (esim. EQ-5D tai 15D). Tilakohtaiset terveydentilaa arvioivat mittarit ovat usein päävastemuuttujina vaikuttavuustutkimuksissa. On tärkeää, että käytetään mittareita, jotka ovat riittävän luotettavia tilan vaikeusasteen määrittelyssä sekä vaikuttavuuden toteamisessa – siis herkkyydessä osoittaa muutos. Sekä yksilön että yhteiskunnan kannalta tärkeä mittari on esimerkiksi kyky osallistua yhteiskunnan toimintaan kuten koulunkäyntiin tai työelämään.

”On tärkeää, että käytetään mittareita, jotka ovat riittävän luotettavia tilan vaikeusasteen määrittelyssä sekä vaikuttavuuden toteamisessa – siis herkkyydessä osoittaa muutos.”

Sosiaalihuollon palveluiden vaikuttavuuden arviointiin on käytettävissä CAIMeR-viitekehys (Blom & Moren, 2010; Matthies ym., 2021; Malmivaara, 2023). Nimi muodostuu englanninkielisistä sanoista *Context, Actors, Intervention, Mechanism, Results*. Konteksteihin (*contexts*) kuuluvat esimerkiksi perus- ja ihmisoikeuksien edistäminen sekä osallisuutta edistävä hyvinvointipolitiikka. Toimijat (*actors*) -käsitteeseen sisältyvät muun muassa asiakkaiden toimijuuden ja siihen vaikuttavien tekijöiden tunnistaminen ja ammatillinen itsereflektointi. Interventioihin kuuluvat muun muassa luottamuksellinen ja kunnioittava yhteistyösuhde. Mekanismit (*mechanisms*) voivat olla sosiaalisia, sosiaalipsykologisia tai psykologisia, ja työmenetelminä esimerkiksi haastaminen ja riskinotto. Tulokset ja vaikutukset (*results and effects*) sisältävät muun muassa sosiaalityön ja monialaisten palvelujen vaikutusten arvioinnin sekä sosiaalityön vaikutuksen kestävään siirtymään.

CAIMeR-viitekehysten periaatteet vaikuttavuuden arvioinnissa ovat samat kuin PICOS-viitekehyksessä, mutta CAIMeR-menetelmässä dokumentoidaan myös sosiaalinen konteksti, jossa vaikuttavuuden arviointi tapahtuu. Lisäksi avun kohteena olevan asiakkaan määrittelyn rinnalle otetaan sosiaalihuollon ammattilainen, ja intervention kuvauksen lisäksi kuvataan oletettu vaikutusmekanismi. Tulostittaus on kontekstisidonnaista kuten PICOS-viitekehystä käytettäessä.

Yksilöön kohdistuvien toimenpiteiden vaikuttavuuden arviointi eli kliininen impaktitutkimus

Kliininen impaktitutkimus (*clinical impact research*) kattaa kaikki tutkimukset, joiden tarkoituksena on arvioida yksilölle suunnattujen sosiaali- ja terveydenhuollon tai kansanterveystyön toimenpiteiden impakteja terveyteen ja hyvinvointiin (Malmivaara, 2016a). Impakteilla tarkoitetaan sosiaali- ja terveydenhuollon päätavoitteita, jotka ovat toimenpiteiden saatavuus ja saavutettavuus, laatu, yhdenvertaisuus, vaikuttavuus, turvallisuus ja kustannusvaikuttavuus. Tutkimuskysymyksen mukaan kliininen impaktitutkimus voi olla satunnaistettu vertailututkimus tai vertaiskontrolloitu tutkimus. Kliinisen impaktitutkimuksen tutkimuskysymys muotoillaan PICOS-viitekehysten mukaisesti.

Kaksoissokkoutettu asetelma – hoidon biologinen vaikuttavuus

Toimenpiteen pelkkää biologista vaikuttavuutta on mahdollista tutkia silloin, kun menetelmä on lääkkeellinen, kohdistuu potilaan anatomiaan ja fysiologiaan (kuten leikkaushoito) tai potilaaseen kohdistetaan sähkömagneettista tai mekaanista energiaa (kuten ultraäänihoidossa). Näiden interventioiden vaikuttavuuden selvittämiseksi on välttämätöntä tehdä kaksoissokkoutettu satunnaistettu kontrolloitu koe (*double blinded randomized controlled trial*), jossa hoitoa antavat henkilöt ja potilas eivät tiedä, onko annettu hoito vaikuttavaa vai pelkkää lumehoitoa (Malmivaara, 2018a). Kaksoissokkoutettua koasetelmaa on järkevää käyttää ainoastaan sellaisissa menetelmissä, joissa on biologinen ydinvaikutus ja tavoitteena on ainoastaan tämän ydinvaikutuksen määrittäminen. Lasten ja

nuorten mielenterveyttä edistävät psykososiaaliset menetelmät eivät täytä näitä edellytyksiä: kaksoissokkoutuksen käyttöön ei ole perusteita, eikä se ole mahdollistakaan. Näiden tutkimusten luotettavuutta arvioitaessa kaksoissokkoutuksen puutetta ei pidä käyttää merkinä tutkimustulosten heikentyneestä luotettavuudesta.

Yksittäisen intervention tai interventiokokonaisuuden vaikuttavuus

Kun tutkitaan yksittäisten menetelmien vaikuttavuutta, luotettavin tutkimusasetelma vaikuttavuuden arvioimiseksi on yleensä satunnaistettu vertailututkimus (Malmivaara, 2016a). Yksittäisten menetelmien vaikuttavuutta voidaan joissakin tapauksissa tutkia myös vertaiskontrolloidulla tutkimuksella, tai tämä voi olla jopa ainoa vaihtoehto tutkimuskysymykseen vastaamiseksi. Vertaiskontrolloidun tutkimuksen aiheena voivat olla eettiset tai tutkimuskysymykseen liittyvät syyt tai satunnaistetun kokeen käytännön toteutuksen vaikeudet (Malmivaara, 2016a).

Mikäli hoidolla arvioidaan kaksoissokkoutetun satunnaistetun vertailututkimuksen tulosten perusteella olevan biologista vaikuttavuutta, tulisi arjen vaikuttavuus määrittää avoimessa satunnaistetussa vertailututkimuksessa. Avoimessa tutkimusasetelmassa sekä hoitavat henkilöt että asiakkaat tietävät, ketkä asiakkaista saavat hoidon, jolla on ydinvaikutus ja ketkä eivät sitä saa (Malmivaara, 2018a). Edellisessä ryhmässä asiakkaat saavat sekä biologisen vaikutuksen että lumevaikutuksen, mutta jälkimmäisessä ryhmässä asiakkaat jäävät sekä biologista vaikutusta että lumevaikutusta paitsi.

”Kun tutkitaan yksittäisten menetelmien vaikuttavuutta, luotettavin tutkimusasetelma vaikuttavuuden arvioimiseksi on yleensä satunnaistettu vertailututkimus.”

Satunnaistetun vertailukoetutkimuksen merkittävä etu on yleensä alkutilanteen vertailukelpoiset ryhmät. Kokeellisessa tutkimuksessa tutkimusryhmien välinen menetelmien ero (ns. interventio-kontrasti) saattaa kuitenkin kaveta heikon hoitoon sitoutumisen vuoksi. Tämä heikentää tutkimusasetelman luotettavuutta. Vertaiskontrolloidun tutkimuksen etuna on interventioiden toteutuminen arjen todellisuuden mukaisesti, mutta haittana satunnaistettua koetta heikompi ryhmien välinen vertailukelpoisuus alkutilanteessa.

Lasten ja nuorten mielenterveyttä edistävillä psykososiaalisilla menetelmillä tavoitellaan vaikuttavuutta fyysisten, psykologisten sekä ympäristöön ja yhteisöihin vaikuttavien toimien avulla. Kun hoito ei perustu pelkkään yksittäiseen biologiseen hoitoon, ei vaikuttavuutta ole mieltä eikä edes mahdollista arvioida kaksoissokkoutetulla satunnaistetulla vertailukokeella, vaan kokeen tulee aina olla avoin.

Hoitoketjun vaikuttavuus

Hoitoketjujen vaikuttavuutta, turvallisuutta ja kustannusvaikuttavuutta selvittäessä vertaiskontrolloitu tutkimus on ensisijainen menetelmä. Tutkimuskysymyksenä on käytössä olevien hoitoketjujen välisten vaikuttavuuserojen selvittäminen. Tutkimuksesta saatava tieto on suoraan sovellettavissa arkeen ja mahdollistaa vertaiskehittämisen eli oppimisen parhaista käytännöistä (*benchmarking*).

”Hoitoketjujen vaikuttavuutta, turvallisuutta ja kustannusvaikuttavuutta selvittäessä vertaiskontrolloitu tutkimus on ensisijainen menetelmä.”

Satunnaistettu vertailukoetutkimus edellyttää koejärjestelyä, jossa palvelujen tuottajia arvotaan noudattamaan heille langennutta, etukäteen määriteltyä hoitoketjua. Tämä ryhmäsatunnaistettu vertailukoeasetelma edellyttää tutkittavien hoitoketjun implementoimista käytäntöön ja antaa vastauksen ainoastaan tutkimusprotokollassa päätettyjen hoitoketjujen paremmuuseroihin. Arjen hoitoketjut eivät yleensä ole samankaltaiset kuin kontrolloidussa vertailukokeessa, mikä heikentää tulosten sovellettavuutta.

Lasten ja nuorten mielenterveyttä edistävä psykososiaalinen menetelmä voi olla osa sosiaali- ja terveydenhuollon hoitoketjua, jolloin sen sisällön ja toteutumisen kuvaaminen on vaikuttavuuden arvioinnin kannalta tärkeää.

Vertaisten välinen vaikuttavuus

Samankaltaisia asiakasryhmiä hoitavien palveluntarjoajien keskinäinen vertaiskehittäminen edellyttää aina vertaiskontrolloitua tutkimusasetelmaa, jonka pohjana on asiakasryhmäkohtainen rekistereistä saatava tai vastaava tieto (Malmivaara, 2022a; Simula ym., 2024). Vertailtavana voi olla lasten ja nuorten mielenterveyttä psykososiaalisin menetelmin edistävät sosiaali- ja terveydenhuollon toimintayksiköt tai muut palvelun tarjoajat.

Vaikuttavuuden arviointi eli järjestelmäimpaktitutkimus

Sosiaali- ja terveydenhuoltoa ja kansanterveystyötä ohjaaviin järjestelmiin sekä terveyspolitiikkoihin kohdistuvat toimenpiteet pyrkivät samaan tavoitteeseen kuin, mihin asiakastyössä pyritään eli tuottamaan parempaa vaikuttavuutta asiakkaille ja väestölle. Tämän vuoksi on tarpeen tutkia järjestelmätason piirteiden vaikuttavuutta ja edistää piirteitä, jotka optimoivat kustannusvaikuttavuuden ja väestön yhdenvertaisuuden. Järjestelmillä ei sinänsä ole vaikuttavuutta, vaan järjestelmät mahdollistavat eriasteisesti sen, missä määrin asiakkaita auttavat ammattilaiset voivat tuottaa terveyttä ja hyvinvointia kullekin yksilölle. Esimerkiksi hyvin organisoidut hoitoketjut tuottavat parempaa vaikuttavuutta kuin heikosti organisoidut.

Sosiaali- ja terveydenhuoltojärjestelmien ja kansanterveystyön tuottama vaikuttavuus toteutuu vaikuttavuuden määritelmän mukaisesti aina yksilötasolla, jolla myös vaikuttavuuden arvioinnin tulee tapahtua. Järjestelmäimpaktitutkimuksilla (*system impact research*) pyritään arvioimaan järjestelmän ominaisuuden tai sen muuttamisen impakteja asiakkaisiin tai väestöön. Impakteilla tarkoitetaan toimenpiteiden saatavuutta ja saavutettavuutta, laatua, yhdenvertaisuutta, vaikuttavuutta, turvallisuutta ja kustannusvaikuttavuutta. Järjestelmien vaikuttavuus, turvallisuus ja kustannusvaikuttavuus ovat osoitettavissa ainoastaan määritellyillä asiakas- ja väestöryhmillä.

Järjestelmäimpaktitutkimuksen kohteina ovat sosiaali- ja terveydenhuollon järjestelmien ominaisuudet, kuten rahoitus, korvaus- ja kannustinjärjestelmät, organisointi, säädökset, käytettävissä olevat voimavarat ja muut järjestelmän ominaisuudet (Malmivaara, 2016b). Kysymyksenasetteluna on, miten järjestelmään tehtävillä muutoksilla ja investoinneilla voidaan edistää asiakkaiden ja

väestön palvelujen saatavuutta ja saavutettavuutta sekä laatua, vaikuttavuutta, turvallisuutta ja kustannusvaikuttavuutta yhdenvertaisella tavalla. Järjestelmiin kohdistuvien menetelmien vaikuttavuutta voidaan tutkia parhaiten vertaiskontrolloidulla tutkimuksilla. Vaihtoehtona ovat ryhmäsatunnaistetut vertailukokeet, jotka kuitenkin vaativat huomattavia voimavaroja ja joissa kokeen toteutettavuus on haasteellista sekä tietoa voi olla vaikea soveltaa arkeen.

”Järjestelmäimpaktitutkimuksen kohteina ovat sosiaali- ja terveydenhuollon järjestelmien ominaisuudet, kuten rahoitus, korvaus- ja kannustinjärjestelmät, organisointi, säädökset, käytettävissä olevat voimavarat ja muut järjestelmän ominaisuudet.”

Lasten ja nuorten mielenterveyttä edistävien psykososiaalisten menetelmien toteutuminen ja vaikuttavuus ovat riippuvaisia sosiaali- ja terveydenhuoltojärjestelmien ja kansanterveyttä edistävien järjestelmien ominaisuuksista. Järjestelmäimpaktitutkimusten tehtävä on selvittää, millaiset järjestelmät parhaiten edistävät lasten ja nuorten psykososiaalisten menetelmien saatavuutta ja saavutettavuutta sekä laatua, yhdenvertaisuutta, vaikuttavuutta, turvallisuutta ja kustannusvaikuttavuutta.

VAIKUTTAVUUSTUTKIMUSTEN LUOTETTAVUUSKRITEERIT

Satunnaistamalla riittävän suuri määrä tutkittavia saadaan vertailtavat tutkimusryhmät ennusteellisten tekijöiden suhteen keskimääräisiltä ominaisuuksiltaan hyvin samankaltaisiksi. Alkutilanteen vertailukelpoisuus on sekä satunnaistetun vertailututkimuksen että vertaiskontrolloidun tutkimuksen ensimmäinen keskeinen luotettavuuskriteeri.

Menetelmien (interventioiden) toteutumisen ero interventio- ja vertailuinterventioryhmän välillä on sekä satunnaistetuissa vertailukokeissa että vertaiskontrolloiduissa tutkimuksissa syytekijä tuloksille ja toinen keskeinen luotettavuuskriteeri. Asiaa arvioidaan selvittämällä, missä määrin suunnitellut menetelmät todella toteutuivat interventio- ja vertailuinterventioryhmissä, kuinka suuri osa asiakkaista siirtyi interventioryhmästä toiseen sekä missä määrin menetelmän- ja vertailumenetelmän tutkittavat saivat lisäinterventioita, joita ei tutkimussuunnitelmaan sisällynyt. Kaikkien toteutuneiden interventioiden ero eli interventiokontrasti on syytekijänä ryhmien välisille eroille tuloksissa eli vaikuttavuudelle.

”Alkutilanteen vertailukelpoisuus on sekä satunnaistetun vertailututkimuksen että vertaiskontrolloidun tutkimuksen ensimmäinen keskeinen luotettavuuskriteeri.”

Tulosten kattava ja luotettava arviointi on sekä satunnaistetun vertailututkimuksen että vertaiskontrolloidun tutkimuksen kolmas keskeinen luotettavuuskriteeri. Tavoitteena on arvioida, kuinka päteviä ensisijaiset ja toissijaiset vastemuuttajat ovat olleet ja kuinka täydellisesti tutkittavat ovat kussakin vertailtavassa ryhmässä seurantaan osallistuneet.

Satunnaistetuissa vertailututkimuksissa hoitoaieanalyysi (*intention-to-treat analysis*, ITT) tarkoittaa sitä, että tutkittavat analysoidaan niissä ryhmissä, joihin heidät on arvottu. Kun tulokset satunnaistetussa vertailututkimuksessa tehdään sen mukaan, miten interventiot ovat kussakin vertailtavassa ryhmässä toteutuneet, ei tarvita muita analyyseja. Esimerkiksi jos interventioryhmässä kaikki asiakkaat ovat saaneet tutkittavan intervention ja vertailuryhmässä tämän on saanut kolmasosa asiakkaista, tulkitaan tutkimuksen osoittama vaikuttavuus interventioissa toteutuneiden erojen pohjalta eikä tehdä johtopäätöksiä intervention vaikuttavuudesta, ikään kuin kukaan vertailuryhmässä ei olisi interventiota saanut. Hoitoaieanalyysia voidaan täydentää kuvaamalla, kuinka hyvin vertailuryhmästä hoitoryhmään vaihtaneet asiakkaat ovat kussakin seurannan vaiheessa toipuneet päästyään interventiosta osallisiksi (Malmivaara, 2018b).

VAIKUTTAVUUSTUTKIMUSTEN TULOSTEN SOVELLETTAVUUS ARKEEN

Satunnaistetut vertailututkimukset pyritään tekemään mahdollisimman täsmällisesti ja yhdenmukaisesti etukäteen määritellyissä olosuhteissa, jonka kaltaisia ei reaali maailmassa ole. Sosiaali- ja terveydenhuollon arjessa toimenpiteen vaikuttavuus vaihtelee asiakkaiden ominaispiirteiden, menetelmien (interventioiden) toteutumisen, työntekijöiden osaamisen ja järjestelmän piirteiden mukaan.

”Satunnaistetut vertailututkimukset pyritään tekemään mahdollisimman täsmällisesti ja yhdenmukaisesti etukäteen määritellyissä olosuhteissa, jonka kaltaisia ei reaali maailmassa ole.”

Satunnaistettujen vertailututkimusten ja systemaattisten katsaus-ten tulosten sovellettavuus arkeen edellyttää alkuperäistutkimusten kattavaa raportointia kahdella tasolla: miten tutkimus on tarkoitettu toteutettavan (tutkimussuunnitelman mukaan) ja miten tutkimus todellisuudessa toteutui (Malmivaara, 2021). Jälkimmäisen osalta tulee dokumentoida ja raportoida tutkittavien valikoituminen, heidän ominaisuutensa, se, miten toimenpiteet ovat toteutuneet ja ovatko asiakkaat saaneet tutkimussuunnitelmaan sisällyttämiä interventioita, sekä kaikki vaikuttavuus- ja haittatulokset.

Vertaiskontrolloitujen tutkimusten tulokset ovat periaatteessa suoraan sovellettavissa kohteisiin, joissa tiedonkeruu on tapahtunut. Parhaimmillaan vertaiskontrolloitujen tutkimusten lähdeaineistona ovat kattavat ja luotettavat valtakunnalliset laaturekisterit, joihin liittyy jatkuva tiedonkeruu (Keinänen ym., 2020). Myös laaturekisterin kaltainen ei-valtakunnallinen dokumentointi voi mahdollistaa vertailun interventioita antavien tahojen välillä (Simula ym., 2024). Riittävä tutkittavien ominaisuuksien, interventioiden sisällön ja vastemuuttajien kuvaus on sovellettavuuden kannalta erittäin tärkeää niin vertaiskontrolloiduissa tutkimuksissa kuin satunnaistetuissa vertailututkimuksissa.

VERTAISKONTROLLOINNIN MENETELMÄ ARVIOINNIN POHJAKSI

Satunnaistettujen vertailututkimusten ja vertaiskontrolloitujen tutkimusten luotettavuuden ja tulosten yleistettävyyden arvioinnissa suositellaan vertaiskontrollonin menetelmää (*benchmarking method*) (Malmivaara, 2022b). Vertaiskontrollonin menetelmä perustuu viiteen luokkaan ja useisiin alaluokkiin, joissa kuvataan tutkimuskysymys ja sen toteutuminen. Arviointi kohdistuu

- 1) tutkittavien valikoitumiseen
- 2) tutkittavien perusominaisuuksiin
- 3) interventioihin
- 4) tuloksiin
- 5) tilastoanalyysiin.

Vertaiskontrollonin menetelmässä vaikuttavuustutkimukset ja toiminnan arkivaikuttavuus arvioidaan samoilla menetelmillä, olipa asetelma satunnaistettu vertailututkimus tai vertaiskontrolloitu tutkimus. Vertaiskontrollonin menetelmä soveltuu myös systemaattisten katsausten arviointiin. Katsauksiin sisältyvien alkuperäistutkimusten kattava kuvaus ja siltä pohjalta tehtävät johtopäätökset ovat tärkein asia systemaattisten katsausten ja meta-analyyysien arvioinnissa. Systemaattisten katsausten ja meta-analyyysien luotettavuuden kannalta on ratkaisevaa, kuinka kattavasti ja huolellisesti katsaukset ovat kuvanneet tutkimuksensa kohteet, tavallisimmin satunnaistettuja vertailututkimuksia. Vertaiskontrollonin menetelmä paljastaa puutteet alkuperäistutkimusten tiedon kattavuudessa, ja systemaattisissa katsauksissa sen avulla päästään kattavaan kuvaukseen alkuperäistutkimusten tuottamasta tiedosta ja tiedon puutteista. Kun tehdään johtopäätöksiä vaikuttavuudesta, tieto tutkimuksesta puuttuvista tiedoista on yhtä tärkeä kuin tieto tutkimuksessa dokumentoiduista tiedoista.

Tutkimuksen arvioinnissa on neljä asiakokonaisuutta.

1. Miten tutkimussuunnitelmassa on määritelty asiakkaat tai tutkittava väestö, menetelmät (interventiot ja vertailuinterventiot) ja vastemuuttajat (*Population, Intervention, Control intervention* ja *Outcome* eli lyhyemmin PICO)?
 - Mitkä olivat tutkimushenkilöiden mukaanotto- ja poissulkukriteerit?
 - Mikä oli interventioiden ja vertailuinterventioiden sisältö (niin yksityiskohtaisesti kuvattuna, että se on toistettavissa)?
 - Mitkä olivat suunnitellut vastemuuttajat ja mikä tai mitkä näistä on nimetty päävastemuuttajiksi?
2. Miten tutkimus on toteutunut (PICO: tutkittava väestö, menetelmät (interventiot ja vertailuinterventiot) ja vastemuuttajat)?
 - Miten hyvin alkuperäistutkimuksessa on kuvattu tutkittavien valikoituminen, tutkittavien ominaisuudet alkutilanteessa, interventioiden ja vertailuinterventioiden sekä lisäinterventioiden toteutuminen ja se, miten kattavasti ja luotettavasti tulokset saatiin mitattua?

3. Mikä on tutkimusten pätevyys (*validity*) ja harhan riski (*risk of bias*)?

Alkuperäistutkimusten tietojen kattava kuvaus (edellä kohdat 1 ja 2) sisältää satunnaistetun vertailututkimuksen ja vertaiskontrolloidun tutkimuksen luotettavuuden arvioinnin keskeiset piirteet. Myös systemaattisen katsauksen sisäinen validiteetti tulee suurelta osalta katettua. Systemaattiselle katsaukselle ominaisia luotettavuusominaisuuksia ovat kuitenkin myös

- kattava alkuperäistutkimusten tavoittaminen (kattavat sähköiset ja muut haut)
- alkuperäistutkimusten huolellinen luokittelu tutkimuskysymyksen mukaan ja meta-analyyysin tekeminen vain niistä tutkimuksista, jotka ovat tutkineet samaa kysymystä (kliininen homogeenisuus)
- kahden tutkijan toisistaan riippumatta tekemä alkuperäistutkimusten valinta ja tiedon dokumentointi
- mahdollisen julkaisuharhan riskin arviointi. Julkaisuharha tarkoittaa sitä, että osa aiheeseen liittyvistä tutkimuksista on jäänyt julkaisematta.

4. Mikä on tutkimusten sovellettavuus?

Mahdollisuus arvioida tutkimustulosten sovellettavuutta riippuu siitä, kuinka tarkasti edellä esitetyn mukaisesti asiakkaat, interventiot ja vastemuuttajat on kuvattu satunnaistetuissa vertailututkimuksissa ja vertaiskontrolloiduissa tutkimuksissa. Sama koskee systemaattisia katsauksia ja meta-analyysejä, joiden kohdalla on myös kyse siitä, kuinka perustellusti (aihealueen tuntevien arvioimana) alkuperäistutkimukset on luokiteltu.

VERTAISKONTROLLOINNIN MENETELMÄ VAIKUTTAVUUSTUTKIMUSTEN ARVIOINTIIN

Vertaiskontrollonin menetelmää voidaan käyttää sekä satunnaistettujen vertailututkimusten että vertaiskontrolloitujen tutkimusten suunnitteluun, toteutukseen, raportointiin ja arviointiin (Malmivaara, 2022b).

Vertaiskontrollonin menetelmää voidaan hyödyntää tutkimuksissa, joissa arvioidaan yhden tai useamman palvelun vaikuttavuutta, hoitopolkujen vaikuttavuutta, vertaisten keskinäistä suoriutumista tai järjestelmiin liittyvien tekijöiden vaikuttavuutta. Vertaiskontrollonin menetelmällä pyritään selvittämään, miten hyvin arviointikohde on kuvattu ja mitkä ovat kuvauksen puutteet. Tiedon puutteiden tiedostaminen on yhtä tärkeää kuin käytettävissä oleva tieto, kun tehdään johtopäätöksiä vaikuttavuudesta ja kustannusvaikuttavuudesta. Tämä on erityisen tärkeää, kun arviointikohde on kompleksinen, kuten lasten ja nuorten mielenterveyttä edistävässä psykososiaalisissa menetelmissä. ●

AVAINSANAT:

vaikuttavuus, arkivaikuttavuus, PICO, satunnaistettu vertailulukko, vertaiskontrolloitu tutkimus

SIDONNAISUUDET:

Kirjoittajalla on hankintasopimus Itlan kanssa. Hän on toiminut Itlan Kasvun tuen vaikuttavuuden arvioinnin asiantuntijana. Artikkelin on kirjoitettu Itlan toimeksiannosta. Artikkelin näkemykset ovat kirjoittajan omia.

LÄHTEET:

- Blom, B. & Moren, S. (2010). Explaining Social Work Practice—The CAIMeR Theory. *Journal of Social Work*, 10(1), 98–119. <https://doi.org/10.1177/1468017309350661>
- Keinänen, M., Brummer-Korvenkontio, H., Eskola, M., Ettala, O., Hartikainen, J., Hovi, P., ... & Malmivaara, A. (2020). Terveystieteiden kansalliset laaturekisteripilotit loppusuoralla. *Suomen Lääkärilehti*, 75(50), 2788–2794. <https://urn.fi/URN:NBN:fi-fe2021041310267>
- Malmivaara, A. (2022a). Vaikuttavuus sosiaali- ja terveydenhuollossa. *Kustannus Oy Duodecim*.
- Malmivaara, A. (2016a). Clinical Impact Research—How to choose experimental or observational intervention study? *Annals of Medicine*, 48(7), 492–495. <https://doi.org/10.1080/07853890.2016.1186828>
- Malmivaara, A. (2016b). System impact research—Increasing public health and health care system performance. *Annals of Medicine*, 48(4), 211–215. <https://doi.org/10.3109/07853890.2016.1155228>
- Malmivaara, A. (2020). Vision and strategy for healthcare: Competence is a necessity. *Journal of rehabilitation medicine*, 52(5), jrm00061. <https://doi.org/10.2340/16501977-2684>
- Malmivaara, A. (2022b). Miten sosiaali- ja terveydenhuollon vaikuttavuutta pitäisi tutkia? *Suomen Lääkärilehti*, 77, 1100–1104.
- Malmivaara, A. (2023). Yhteistä säveltä sosiaali- ja terveydenhuollon arkivaikuttavuuden arviointiin. *Yhteiskuntapolitiikka*, 88(4), 430–438. <https://urn.fi/URN:NBN:fi-fe20230913124929>
- Malmivaara, A. (2021). Applicability of evidence from randomized controlled trials and systematic reviews to clinical practice: A conceptual review. *Journal of Rehabilitation Medicine*, 53(6) 2800. <https://doi.org/10.2340/16501977-2843>
- Malmivaara, A. (2015). Benchmarking Controlled Trial—a novel concept covering all observational effectiveness studies. *Annals of Medicine*, 47(4), 332–340. <https://doi.org/10.3109/07853890.2015.1027255>
- Malmivaara, A. (2018a). Pure intervention effect or effect in routine health care—Blinded or non-blinded randomized controlled trial. *BMC Medical Research Methodology*, 18(1), 91. <https://doi.org/10.1186/s12874-018-0549-z>
- Malmivaara, A. (2018b). Validity and generalizability of findings of randomized controlled trials on arthroscopic partial meniscectomy of the knee. *Scandinavian Journal of Medicine & Science in Sports*, 28(9), 1970–1981. <https://doi.org/10.1111/sms.13215>
- Malmivaara, A., Zampolini, M., Stam, H. & Gutenbrunner, C. (2022). Pros and Cons of Randomized Controlled Trials and Benchmarking Controlled Trials in Rehabilitation: An Academic Debate within the European Academy of Rehabilitation Medicine. *Journal of Rehabilitation Medicine*, 54, jrm00319. <https://doi.org/10.2340/jrm.v54.2511>
- Matthies, A.-L., Svenlin, A.-R. & Turtiainen, K. (2021). Aikuissosiaalityö. Tieto, käytäntö ja vaikuttavuus. *Gaudeamus*.
- Miettinen, O. S., Steurer, J. & Hofman, A. (2019). *Clinical research transformed*. Springer, Cham, Journal Article.
- Miettinen, O. S. (2020). *Henkilökohtainen tiedonanto*.
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gotzsche, P. C., Devereaux, P. J., ... & Altman, D. G. (2010). CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. *BMJ*, 340, c869–c869. <https://doi.org/10.1136/bmj.c869>
- Morabia, A. (2006). Pierre-Charles-Alexandre Louis and the Evaluation of Bloodletting. *Journal of the Royal Society of Medicine*, 99(3), 158–160. <https://doi.org/10.1177/014107680609900322>
- Simula, A. S., Malmivaara, A., Booth, N. & Karppinen, J. (2024). Effectiveness of a classification-based approach to low back pain in primary care – a benchmarking controlled trial: Trust your back. *Journal of Rehabilitation Medicine*, 56, jrm28321. <https://doi.org/10.2340/jrm.v56.28321>