

Vaikuttavuustutkimuksen pulmallisuus psykoterapiassa

Mikael Leiman

Hans Eysenckin »meta-analyysi«vuodelta 1952 **14** Kultainen standardi psykoterapia-tutkimuksessa **14** Vanderbilt I **15** Ekvivalenssiparadoksi **16** Yritetään vielä kerran **17** Terapiamuotoa olennaisempia vaihtelun lähteitä **17** Yritetään uudestaan **19** Psykoterapia kehitystapahtumana **19** Yhteistoiminta on vaikuttava ainesosa **20** Vaihtoehtoisia metodisia otteita **20**

Näyttöön perustuva lääketiede on viime vuosina yleistynyt ajattelutapa hoidon tutkimisen alueella. Psykoterapia muodostaa tämän ajattelutavan kannalta kiintoisan ja ongelmallisen tutkimushaasteen. Kirjoituksessa tarkastellaan psykoterapian tuloksellisuustutkimuksen metodisia pulmia ja ratkaisutapoja tutkimuksen 50-vuotisen historian ajalta. Tarkastelun valossa kokeellisten ryhmävertailujen yksinomainen käyttö näytön lähteenä on pulmallista. Kirjoituksessa esitellään vaihtoehtoisia tutkimusotteita, joita psykoterapian tutkimuksen piirissä on kehitetty kokeellisten asetelmien synnyttämien ongelmien ratkaisemiseksi.

Terveystieteiden rahoitusjärjestelmien muutokset sekä terveystaloustieteen ja vakuutuslääketieteen kehitys ovat tehneet hoitojen tuloksellisuuden tutkimuksesta erittäin kysytyn tutkimusalan. Näyttöön perustuvan lääketieteen ajattelutavan yleistyessä myös psykoterapioiden tuloksellisuuden osoittamista eri tahoilla alettu korostaa. Suomalainen Lääkäriseura Duodecim on hiljattain julkaissut *Käypä hoito* -suosituksen depression hoidosta. Siinä hoitomuotoja koskevat suositukset perustuvat tämän ajattelutavan mukaisesti kokeellisille ryhmävertailuille ja niistä laadittuihin meta-analyysiin.

Psykoterapioilla on kuitenkin sellaisia ominaisuuksia, jotka tekevät niistä erittäin hankalia tutkia kokeellisen ryhmävertailun menetelmällä. Hoidon sisältö muodostuu psykoterapeutin ja

potilaan kommunikoinnista. Sille on ominaista, että kummankin toiminta muotoutuu vastauksena toiselle. Tämä psykoterapioille luonteenomainen vastavuoroisuus on ristiriidassa kokeellisissa tutkimuksissa sovellettavan metodologian perusoletusten kanssa.

Kun lääketutkimuksesta peräisin olevia kokeellisia menettelytapoja on yritetty soveltaa psykoterapioihin, tulokset ovat olleet epämääräisiä ja ristiriitaisia. Yli viidenkymmenen vuoden mittainen historia tarjoaa kiehtovan näköalan tämän hoitomuodon tutkimuksen metodologisiin ongelmiin. Niiden ymmärtäminen hyödyttää kaikkia sellaisia hoidon tutkimuksen aloja, joilla hoitovaikutus perustuu hoidettavan ja hoitajan yhteistoimintaan ja vuorovaikutukseen.

Hans Eysenckin »meta-analyysi»vuodelta 1952

Psykoterapian järjestelmällisen tuloksellisuustutkimuksensiiivittäjänä pidetään brittiläisen psykologin Hans Eysenckin v. 1952 julkaisemaa artikkelia *The effects of psychotherapy*. Eysenckin välittömänä tavoitteena oli torjua American Psychological Associationin (APA) piirissä syntynyt ehdotus psykoterapian sisällyttämisestä kliinisen psykologian jatkokoulutukseen. Hän halusi osoittaa, että vaikutuksiltaan olemattoman hoitomuodon koulutuksessa ei ole mitään mieltä. Tavoitteensa Eysenck sai läpi lähinnä vain omassa kotimaassaan Englannissa, jossa kliinisen psykologian opetukseen ei sisällynyt vuosikymmeniin psykoterapiaa. Artikkelilla oli kuitenkin huomattava merkitys kokeellisiin ryhmävertailuihin perustuvan psyko-terapian tutkimuksen kehityksessä.

Eysenckin menetelmää voidaan pitää vapaa- muotoisena meta-analyysinä siinä mielessä, että hän teki päätelmänsä psykoterapioiden tehottomuudesta olemassa olevaa tutkimustietoa tiivistämällä. Artikkelin argumentit ovat tosin kestävämpiä meta-analyysissä nykyisin sovellettavien tilastollisten kriteerien valossa, mutta puoli vuosisataa sitten ei tällaisia metodisia sääntöjä vielä ollut.

Artikkelin pohjana oli 19 julkaistua psykoterapian seurantatutkimusta. Viidessä tarkasteltiin psykoanalyysiä ja 14:ssä eri suuntauksia yhdisteiviä hoitoja. Tutkimukset perustuivat yksinkertaisesti arviointeihin potilaan tilasta hoitojen päätyttyä ja tietyn seuranta-ajan jälkeen. Koska vertailuryhmiä ei ollut, Eysenck otti vertailupohjaksi kaksi tutkimusta, joissa oli tarkasteltu neuroottistaisoisten potilaiden vointia 1–2 vuoden kuluttua siitä, kun he olivat ongelmansa vuoksi olleet yleislääkärin hoidossa tai valtion mielisairaalassa. Sekä terapia- että vertailututkimusten aineiston hän ryhmitteli uudelleen neliluokkaisen tuloksellisuusmuuttujan perusteella.

Koska yleislääkärien hoitamien ja mielisairaloissa hoidettujen potilaiden paranemistulokset olivat neliluokkaisella arviointiasteikolla vähintään yhtä hyvät tai paremmat kuin varsinaisten

psykoterapioiden tulokset, Eysenck päätteli, ettei psykoterapiakoulutusta missään nimessä pitäisi sisällyttää lääkärin ja psykologien ammatilliseen koulutukseen.

Eysenckin artikkeli herätti keskustelua ja kritiikkiä. Metodisesti oli selvää, että vertailuryhmää ei voi rakentaa koeryhmistä kokonaan erillisten tutkimusten aineistoista. Myös neliluokkaista tuloksellisuusmuuttujaa voitiin perustellusti pitää erittäin puutteellisena. Artikkelin tärkein anti olikin metodologisen keskustelun viriämisessä.

Kultainen standardi psykoterapiatutkimuksessa

Eysenckin artikkelilla oli stimuloiva vaikutus psykoterapian tuloksellisuustutkimuksen kehitykseen. Tutkimusten lukumäärä lisääntyi nopeasti, ja niiden metodisia periaatteita selventävä keskustelu oli vilkasta. Koska psykoterapiat ovat mielen-terveysongelmien käsittelytapoja, ne rinnastettiin sairauksien hoitoon, ja niiden vaikuttavuuden tutkimisen metodologia lainattiin asiaa suuremmin pohtimatta lääketieteen puolelta. Eysenckin kritiikin herättämässä keskustelussa tavoitteeksi asetettiin psykoterapioiden »vaikuttavan ainesosan» erottaminen spontaanista paranemisesta ja hoitojen suggestiovaikutuksista. Kysymyksenasettelulla oli myös vahva sosiaalinen motiivi: varsinkin Yhdysvalloissa psykoterapiat olivat yksityistä liiketoimintaa, ja näytön hankkiminen tietyn terapiamuodon tehokkuudesta vaikutti heti sen markkina-asemaan ja -potentiaaliin.

Kokeellisesta ryhmävertailusta tuli psykoterapiatutkimuksen metodinen ihanne. Käytännössä sen toteuttaminen oli kuitenkin vaikeaa. Varhaisia käyttäytymisterapioita lukuun ottamatta psykoterapiat kestivät kauan, 2–5 vuotta. Tutkiminen oli siten erittäin kallista. Homogeenisten potilasryhmien kokoaminen oli vaikeaa osaksi diagnostiikan kehittymättömyyden vuoksi, osaksi tietyn hoitomuodon vähäisen saatavuuden johdosta. Myös vertailuryhmien määrittäminen tuotti omat ongelmansa. Apua hakevat ihmiset pyrkivät tavalla tai toisella saamaan apua, eikä

lumelääkkeen kaltaista vertailuhoitoa psykoterapialle voitu kuvitellakaan.

Vaikeuksista huolimatta seuraavan vuosikymmenen aikana psykoterapian tuloksellisuutta tutkittiin melko paljon – eniten käyttäytymisterapioiden ja mielenterveydellisen neuvonnan piirissä, koska niitä oli helpointa tutkia.

1960-luvun loppuvuodet olivat tuon ensimmäisen tutkimusaallon arvioinnin aikaa. Vuonna 1971 ilmestyi Allen Berginin ja Sol Garfieldin toimittama *Handbook of psychotherapy and behavioral change*, johon tutkimuksen tuottamat kokemukset kiteytettiin. Psykoterapiat olivat tuloksellisia monella tavalla arvioituna – myös sellaisten tutkimusten mukaan, joissa hoitoa saaneita ja sitä vaille jääneitä oli voitu verrata. Kysymys eri hoitomuotojen paremmuudesta oli sen sijaan pulmallisempi. Kun eri terapiamuotoja verrattiin toisiinsa, tulokset eivät näyttäneet eroavan merkittävästi toisistaan.

Bergin ja Garfield pohtivat tätä ongelmaa ja tekivät tärkeän havainnon: terapiamuodosta riippumatta hoitoryhmien sisäinen vaihtelu lisääntyi tuloksellisuutta kuvaavilla muuttujilla tutkittuna enemmän kuin vertailuryhmien. Hoitaminen sinänsä näytti sekä parantavan että huonontavan potilaita. Kun hoitoryhmien jakautuma laajeni kummastakin päästään, oli ymmärrettävää, että pelkillä keskiarvovertailuilla ei saatu merkitseviä eroja sen enempää eri hoitomuotojen kuin hoito- ja vertailuryhmien välille.

Vanderbilt I

Hyvään laatuun pyrkivissä tuloksellisuustutkimuksissa oli 1970-luvulla monta ongelmaa ratkottavana. Koeryhmien sisäinen vaihtelu eli potilaiden ongelmien yksilöllisyys oli yksi. Toinen pulma oli, miten varmistaa, että koeryhmille tarjottava psykoterapia oli laadultaan samanlaista. Terapeutit ovat yksilöitä kuten potilaatkin. Heidän tapansa »annostella vaikuttava aines» saattoi vaihdella niin paljon, että yleistetyt päätelmät tietyn terapiamuodon tehosta olivat vailla perustetta. Kolmas ongelma koski vertailuryhmää. Tavan-

Ydinasiat

- Psykoterapia on terapeutin ja potilaan vastavuoroista toimintaa.
- Psykoterapian kannalta on oleellista potilaan suhtautuminen oireeseensa tai ongelmaansa.
- Psykoterapiamuotoa kuvaavat muuttujat hukkuvat aineiston kokonaisvaihteluun.
- Onnistunut psykoterapia muuntaa alkuongelman muotoa ja paikkaa tavalla, jota oiremittarit eivät tavoita.

mukainen hoito tai sijoittaminen jonoon kokeen ajaksi eivät olleet niin neutraaleja toimenpiteitä kuin kokeellisen logiikan toimimisen kannalta olisi suotavaa. Ihmiset etsivät apua ongelmiinsa monin tavoin, vaikka eivät saisikaan tutkimuksen tarjoamaa psykoterapiaa. Vertailuryhmän sisäinen vaihtelu ei suinkaan ollut satunnaista.

Hans Struppin 1970-luvulla johtama Vanderbilt I -projekti (Strupp ja Hadley 1979) suunniteltiin asetelmaltaan sellaiseksi, että näihin kysymyksiin voitiin vastata aiempaa tyydyttävämmin. Potilasryhmän vaihtelua pyrittiin vähentämään ottamalla tutkimukseen vain nuoria miesopiskelijoita, joilla oli ahdistuneisuusongelmia. Ryhmän heterogeenisuutta pyrittiin vielä vähentämään käyttämällä MMPI-testiä (Minnesota Multiphasic Personality Inventory) siten, että vain tietyn testiprofilin saaneet otettiin mukaan aineistoon. Hoitoryhmän terapeutit olivat kaikki hyvin kokeneita ja psykodynaamisesti suuntautuneita. Heillä oli työkokemusta keskimäärin 23 vuotta. Vertailtavaksi interventioiksi valittiin maallikkokeskustelu. Yliopistonopettajat tapasivat opiskelijoita sovitun ajan ja keskustelivat heidän ongelmistaan. Ryhmien sisäisen vaihtelun tutkimiseksi sekä terapeuttien interventiot että opettajien keskustelut nauhoitettiin. Näiden koeryhmien lisäksi asetelmaan kuului kaksi verrokkiryhmää. Toinen oli viivästetyn hoidon ryhmä ja toisen ryhmän jäsenillä oli todettu samanlainen MMPI-profiili, mutta heille ei tehty mitään toimenpidettä.

Huolellisesti toteutetun ryhmävertailun tulos oli Struppille järkytys. Tulosmuuttujissa ei näkynyt merkitsevää eroa vertailu- ja hoitoryhmän välillä.

Yliopiston opettajat olivat tilastollisten tunnuslukujen valossa yhtä tehokkaita sosiaalisen ahdistuneisuuden hoitamisessa kuin pitkään koulutetut psykoterapeutit.

Tuloksia raportoidessaan Strupp kirjoitti kieläytyvänsä uskomasta, että yli kahdenkymmenen vuoden ammattikokemuksella ei olisi mitään merkitystä auttamisen laadun kannalta. Koska kaikki interventiot oli nauhoitettu, oli mahdollista ryhtyä selvittämään, mitä koe- ja vertailuryhmän sisäiseen vaihteluun kätkeytyi.

Potilasryhmän heterogeenisuus osoittautui edelleen ongelmalliseksi. Diagnostisista tarkennusyrityksistä huolimatta nuorten miesten sosiaalinen ahdistuneisuus oli varsin moniulotteista. Kuvailevalla diagnostiikalla, jota psykiatrinen tautiluokittelukin edustaa, ei tavoiteta oireiden taustalla olevia ilmiöitä. Ne liittyvät persoonallisuuden kehitykseen ja yksilöhistoriaan. Diagnostisesta yhdenmukaisuudesta huolimatta aineistossa esiintyi paljon vaihtelua oireilun yksilöllisessä taustassa. Kumpaankin ryhmään sijoittui sekä suhteellisen selväpiirteisiä että mutkikkaita tapauksia, jotka nykykäytännön mukaan saisivat todennäköisesti kaksoisdiagnoosin.

Toinen vaihtelua tuottava seikka on potilaan suhtautuminen oireiluunsa. Tämä tulee näkyviin vasta, kun häntä ryhdytään hoitamaan. Psykoterapeutin tehtävät ovat melko erilaisia sen mukaan, pyrkiikö potilas käsittelemään ahdistustaan kiistämällä, vetäytymällä, avuttomuudella vai uteliaisuudella. Yksittäisiä terapiaprosesseja tutkiessaan Strupp työtovereineen havaitsi, että juuri näiden suhtautumisilmiöiden käsittelyssä ammattiterapeutit työskentelivät huomattavasti vivahteikkaammin kuin yliopisto-opettajat. Ero oli selvin kyvyssä käsitellä potilaiden negatiivisia ilmaisuja ja suhtautumistapoja.

Vaihtelua aineistoihin toivat myös terapeuttien ja opettajien yksilölliset toimintatavat. Kullakin oli tapansa ja tyyliensä. Jos ne olisivat piintyneitä, terapeutteja voitaisiin ehkä valita persoonallisuustestien perusteella ja siten vähentää ryhmänsäistä vaihtelua. Näin ei valitettavasti ollut. Sekä terapeuttien että opettajien yksilölliset toimintatavat vaihtelivat sen mukaan, keitä he hoitivat. Yksilöllisyys kietoutuu hoitosuhteen vuorovaikutukseen. Terapeutin toimintaa ei todellakaan voida

rinnastaa lääketabletin vaikuttavaan aineeseen, joka sekään ei vaikuta mekaanisesti potilaasta riippumatta.

Ekvivalenssiparadoksi

Vanderbilt-projektin ohella useissa muissakin psykoterapiamuotoja vertailevissa tutkimuksissa saatiin samantapaisia tuloksia. Eri terapiamuotojen välille ei saatu merkitseviä eroja tulosmuuttujien keskiarvovertailuissa. Kun eri psykoterapiamuotojen prosessia on tutkittu tarkemmin, on kuitenkin havaittu, että eri suuntausten terapeutit todella toimivat potilaidensa kanssa hyvin eri tavoin. Hoitomuodot ovat siis käytännössä erilaisia, mutta ryhmävertailuun perustuvissa tutkimusasetelmissä tekniikkamuuttujat näyttävät hukkuvan aineiston kokonaisvaihteluun. Ilmiötä ryhdyttiin kutsumaan 1980-luvulla ekvivalenssiparadoksiksi: teetpä niin tai näin, lopputulos on sama.

Hämmäntävään ongelmaan esitettiin kolmenlaisia vastauksia. Ensimmäiseksi tarjoutui pyrkimys pienentää koeryhmien sisäistä vaihtelua. Jos ryhmävertailumetodiikasta halutaan edelleen pitää kiinni, on aineiston laatua parannettava tilastolliselta kannalta. Siis enemmän samaa, mutta paremmin toteutettuna. Toinen mahdollisuus oli pyrkiä ymmärtämään tuon vaihtelun laatua. Entäpä jos psykoterapioille yhteisiä tekijöitä onkin enemmän kuin niitä erottelevia? Carl Rogersin työstä 1950-luvun lopulla alkaneessa neuvonta-interventioiden tutkimuksessa tällaisesta oli jo saatu viitteitä. Kolmas vaihtoehto kyseenalaisti kokeelliset ryhmävertailut ylipäätään järkevänä menettelytapana. Lääketutkimuksen analogia sisältää sellaisia metodologisia oletuksia, jotka eivät vastaa psykoterapiaprosessin luonnetta. Selkeitä tuloksia voi tuskin syntyä, jos niitä etsitään tutkimuskohteen kannalta irrelevantilla välineistöllä.

1980-luvun alussa kaikki nämä vaihtoehdot olivat esillä psykoterapian tutkijoiden kansainvälisen yhteisön *Society for Psychotherapy Researchin* kongresseissa. Kaikkia vaihtoehtoja ryhdyttiin myös selvittämään.

Yritetään vielä kerran

Eikö todellakaan ole mahdollista rakentaa asetelmaa, jossa satunnaisvaihtelu (vielä noihin aikoihin kliiniseltä kannalta epäolennaiseksi oletettu aineiston vaihtelu) saadaan kuriin? Metodinen välineistö oli toki jo paljon hienojakoisempaa kuin kymmenen – saati sitten kolmekymmentä vuotta aiemmin.

Tuoreen mahdollisuuden interventoiden vertailuun tarjosivat jo 1970-luvulla kehittyneet terapiamuodot, kuten interpersonaalinen terapia (IPT) ja kognitiivisbehavioraalinen terapia (CBT), joissa terapeutin toimintaa jäsennettiin ja suunnattiin käsikirjojen avulla. Jos koe- ja verrokkiryhmät muodostettaisiin kattavan satunnaistamisen avulla ja jos hoitoryhmien sisäistä vaihtelua vähennettäisiin manualisoinnin keinoin, ekvivalenssiparadoksi kenties voitettaisiin. Tällöin ei jouduttaisi tinkimään liikaa yleistettävyydestä ja hoidon pienemmätkin tehot saataisiin esille. Tutkimuksen toteuttaminen olisi haastavaa, koska terapeutit pitäisi koulia jokaiseen koeryhmään ja käsikirjan suosituksista poikkeavat hoidot pitäisi karsia aineistosta. Tarvittavan tilastollisen voiman saavuttamiseksi pitäisi hoitojen määrän kussakin koeryhmässä kuitenkin olla mieluummin yli kuin alle sata.

National Institute of Mental Healthin (NIMH) rahoittama suuri depressioprojekti, joka käynnistettiin 1980-luvun alkuvuosina Irene Elkinin johdolla, sisälsi nämä kunnianhimoiset tavoitteet. Tutkimusasetelma rakennettiin erityisen huolellisesti. Hanke toteutettiin monikeskustutkimuksena riittävän tapausmäärän varmistamiseksi. Varsinaiseksi psykoterapiainterventioiksi valittiin IPT ja CBT. Tulosindikaattoreiksi valittiin psykometriasilta ominaisuuksiltaan hyväksi osoittautuneita mittareita, kuten Beckin depressioinventori (BDI) ja Hamilton rating scale for depression (HRSD). Psykoterapian tehokkuuden osoittamiseksi vertailuryhmiä oli kaksi – imipramiinilääkitys ja lumelääkeinterventio tavanomaisine hoito- ja seurantakäynteineen.

Suuri ja kallis ponnistus oli tuloksiltaan ristiriitainen. Sekä IPT että CBT vähensivät merkittävästi oiremittareiden lukemia, mutta hoitojen

keskinäiset erot ja niiden ero imipramiiniryhmään nähden eivät olleet merkitseviä. Ekvivalenssiparadoksi ei siis kumoutunut.

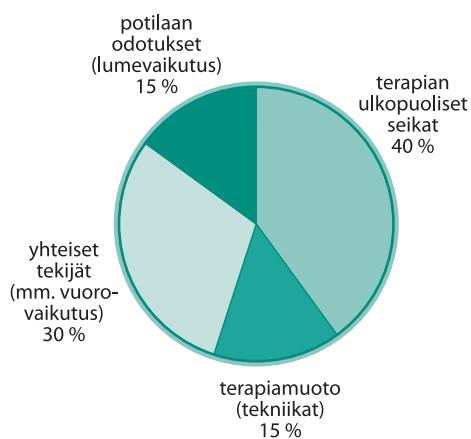
Elkin joutui toteamaan kuten Strupp kymmenen vuotta aiemmin, että psykoterapian salaisuus taitaa piillä aineiston sisällöllisessä vaihtelussa, jota tunnusluvut eivät tavoita. Hyvin suunnitelluissa terapiatutkimuksissa on jo kauan ollut tapana dokumentoida hoitoprosessi nauhoitusten avulla. Näin tehtiin myös NIMH:n tutkimuksessa, ja se on myöhemmin tehnyt mahdolliseksi tutkia eri hoitoryhmien sisäistä vaihtelua samaan tapaan kuin Vanderbilt I -projektissa.

Terapiamuotoa olennaisempia vaihtelun lähteitä

Psykoterapian tutkimuksen toinen suuri linja on painottanut eri terapiamuotojen taustalla olevia yhteisiä tekijöitä. Koska sen kohteena ovat terapia-prosessin ilmiöt, ei kokeellinen ryhmävertailu ole tutkimusasetelmissa olennainen lähtökohta. Tuloksellisuustutkimusten aineistoja on kuitenkin voitu käyttää hyväksi, jos hoidot on nauhoitettu. Niistä on ollut suuri hyöty myös siinä mielessä, että eri ryhmien sisällä on voitu valita tarkemman analyysin kohteeksi tulosmuuttujien mukaan hyvin ja huonosti toteutuneita hoitoja.

Yhteisten tekijöiden tutkimusperinne ulottuu 1950-lopulle, jolloin Carl Rogersin asiakas-keskeiseen neuvontamenetelmään kohdistetuissa tutkimuksissa havaittiin muutaman prosessimuuttujan korreloivan merkittävästi neuvonnan tuloksellisuusmuuttujiin. Terapeutin empaattisuus ja myönteinen yhteistyösuhde osoittautuivat aikaa myöten tekijöiksi, joita kuvaavat muuttujat nousivat johdonmukaisesti esille monimuuttujanalyysissä hoitojen tuloksellisuutta parhaiten ennustavina tekijöinä.

Ekvivalenssiparadoksia valottavaa tietoa eri terapioiden yhteisistä tekijöistä on koottu 1980-luvulta nykypäiviin. American Psychological Association julkisti vuonna 2002 asiantuntijaryhmän raportin, johon on koottu tämän tutkimussuunnan tärkeimmät tulokset (kuva 1).



Kuva 1. Eri muuttujaryhmien prosenttiosuudet psykoterapian tuloksellisuudessa.

Kun tutkitaan eri terapiamuotojen prosessia, käy ilmeiseksi, että terapeutit toimivat eri tavoin terapiatekniikan mukaan. Tekniikkamuuttujien painoarvo on aineiston kokonaisvaihtelussa kuitenkin vain noin 15 %. Valtaosa vaihtelusta on muiden tekijöiden tuottamaa. Tärkeimpiä niistä ovat terapeutin vuorovaikutuksen laatua kuvaavat muuttujat, kuten terapeutin yhteistyösuhteen tunnesävy, terapeutin empaattisuus ja johdonmukaisuus sekä yhteisesti jaettu käsitys hoidon päämääristä. Ne selittävät aineiston vaihtelusta kolmanneksen. Hoitoa tukevat tai ehkäisevät elämänolosuhteet ja -muutokset sekä hoitoon kohdistuvat odotukset tuottavat aineiston vaihtelusta noin 55 %.

APA:n raportissa todetaan, ettei kokeellisten ryhmävertailujen tekemiseen oikeastaan ole perusteita, koska terapeutin vuorovaikutuksen laatu on erityistekniikoita olennaisempi vaihtelun lähde. Ei siis ole ihme, että eri terapiamuotojen vertailussa ei ole saatu eroja näkyviin. Viidentoista prosentin osuus tilastollisesta kokonaisvaihtelusta on yksinkertaisesti liian vähäinen.

Raportissa on esitelty tutkimuksia, joissa on selvitetty, miten manualisoitujen terapioiden edellyttämä terapeutin terapeuttinen kuuliaisuus heijastuu yhteistyösuhteeseen. Tiukasti yleisten suositusten mukaan toimivat terapeutit saavat huonompia tuloksia, koska he eivät jousta yhteis-

työsuhteen kulloistenkin tarpeiden ja edellytysten mukaisesti. Tekniikkamuuttujia korostavat koejärjestelyt tuottavat siten tutkimuksen tavoitteiden kannalta epäsuotavaa vaihtelua aineistoon.

Tutkimusmetodiikka säätelee yhteistyö-suhdetta ja sen laatua. Metodiikka tuottaa muitakin ongelmia. Pyrkinessään kaikin keinoin pienentämään hoitoryhmien potilaskohtaista vaihtelua tutkijat joutuvat suorittamaan jo valikoitumisvaiheessa ankaraa karsintaa, mikä vähentää tulosten yleistettävyyttä ja erityisesti niiden kliinistä relevanssia. Westen ja Morrison havaitsivat depression, paniikkihäiriön ja yleistyneen ahdistuksen hoitoon kohdistuvassa meta-analyysissään, että keskimäärin vain kolmannes potilaista, jotka ongelmansa diagnostisen luokan perusteella olisivat kuuluneet hoidettavaan ryhmään, valittiin lopulta mukaan koeryhmiin. Vertailut, jotka sulkevat valtaosan hoitoon soveltuvista potilaista tutkimuksen ulkopuolelle, antavat kaunistellun kuvan hoidon tuloksellisuudesta.

Kokeellinen tutkimus on altis tutkijoiden intresseille. Luborski ym. teki 29 kokeelliseen vertailuun kohdistuvan meta-analyysin. Siihen valittiin vain erittäin laadukkaita tutkimuksia, jotka arvioitiin kahdentoista metodisen kriteerin mukaisesti. Ryhmä selvitti myös tutkijoiden omaa terapeutista suuntautumista kolmella tavalla: tutkijoiden omaa kantaa kysyttiin, heidät hyvin tuntevilta tutkijatovereilta pyydettiin arvio ja kolmantena menetelmänä oli julkaisujen perusteella tehty arviointi. Nämä arvioinnit korreloitiin hoito- muotojen keskinäistä paremmuutta osoittaviin vaikutusarvoihin. Tutkijan oma suuntautuneisuus ja tutkijan kannattaman hoidon paremmuus vertailtavaan hoitoon nähden korreloivat erittäin voimakkaasti ($r = 0,85$). Regressioanalyysissä tutkijan suuntautuneisuus selitti 69 % hoidon paremmuutta koskevien vaikutusarvojen vaihtelusta.

Vertailut, jotka sulkevat valtaosan hoitoon soveltuvista potilaista tutkimuksen ulkopuolelle, antavat kaunistellun kuvan hoidon tuloksellisuudesta.

Yritetään uudestaan

Psykoterapian tutkimuksen 50-vuotinen historia on osoittanut, että kokeelliset ryhmävertailut ovat monin tavoin ongelmallisia hoidon tuloksellisuutta koskevan näytön lähteinä. 1990-luvun puolivälissä ne kuitenkin tulivat uudestaan käyttöön. Syy tähän oli näyttöön perustuvan lääketieteen voimaperäinen kehitys Yhdysvalloissa ja Britanniassa. Society for Psychotherapy Researchin vuoden 1999 kongressissa vaikutti siltä kuin koko kolmikymmenvuotinen metodologinen keskustelu ja kritiikki olisivat äkillisesti pyyhkiytyneet pois. »Terveyspalvelujen ostajat edellyttävät kokeellista näyttöä, meidän on siis tuotettava sitä.»

Tutkijat ovat riippuvaisia tutkimusten rahoittajien intresseistä. Kun vakuutuslaitokset, Managed Care -yhtiöt ja lopulta myös julkinen terveyshallinto ilmaisevat haluavansa rakentaa toimintansa empiirisesti validoitujen hoitojen pohjalle ja kun ne pitävät lääketutkimuksen metodologiaa ainoana pätevän näytöntuottamistapana, tutkijoiden ja yliopistojen tutkimusyksiköiden on otettava lusikka kauniiseen käteen.

Niin on sitten käynyt, että psykoterapia-alan kansainväliset avainjulkaisut ovat 2000-luvulla täyttyneet ryhmävertailujen tutkimusraporteista. Erityisen ongelmalliseksi tilanteen tekee se, että tutkimuskohteen luonteeseen huonosti sopiva metodinen ote tuottaa tietoa, johon tilaajien ja tutkijoiden intressit pääsevät vaikuttamaan jокseenkin vapaasti.

Psykoterapia kehitystapahtumana

Ryhmävertailujen metodiikka ei kuitenkaan ole ainoa oikea tapa pyrkiä selvittämään psykoterapian vaikuttavuutta. Kolmas, 1980-luvulta lähtien kehittynyt tutkimussuunta hahmottaa psykoterapian kehitystapahtumana. Psykoterapiat eroavat lääketieteellisistä interventioista olennaisesti tuloksellisuuden sisältöön nähden. Somaattisessa hoidossa hyvä tulos on usein määriteltävissä elin-

toimintojen palautumiseksi normaalin vaihtelun alueelle. Homeostaasin periaate on käypä ruumiillisten toimintojen kuvaamiseen, mutta ihmiselämän ja psyykkisten toimintojen kuvaamisessa se on harhaanjohtava.

Psyykkisille muutoksille on ominaista kehityksellisyys: psykoterapiassa tapahtuvan kehityksen lopputulos on ongelmallisen lähtötilanteen muodon muuttuminen. Tämä on keskeinen syy sille, että kokeellisten ryhmävertailujen metodikalla toteutettujen tutkimusten tulokset ovat epämääräisiä ja johdonmukaisin tulos näyttää olevan aineiston vaihtelun kasvu. Samojen mittareiden käyttö alku- ja loppumittauksissa on itse asiassa virheellinen menettelytapa, koska koehenkilöiden tekemien arviointien perusteet ovat muuttuneet hoidon tuloksena. Loppumittaus samoilla mittareilla tuottaa kahdella tavalla irrelevantin tuotoksen. Se pakottaa koehenkilön käyttämään näennäisesti samoja perusteita arvioinneissaan, vaikka ne tosiasiaissa ovat merkityssisällöltään muuttuneet. Se myös jättää arviointien ulkopuolelle ne uudet ilmiöt, jotka terapiassa tapahtunut kehitys on saanut aikaan.

NIMH:n depressioprojektin muuttuja-aineiston uudelleen analysointi muutamia vuosia sitten toi tämän ilmiön varsin selvänä esiin. Yksikään tutkimuksen ns. objektiivisista tulosmuuttujista, kuten BDI ja HRSD tai kliinisten arvioitsijoiden tekemät arviot, eivät erotelleet psykoterapia- ja lääkehoitoryhmiä sen enempää hoitojen lopussa kuin 18 kuukauden seurannassa. Sen sijaan potilaiden tekemät itsearviointit seurantamittauksissa erottelivat psykoterapiaryhmät erittäin merkittävästi imipramiini- ja lumeryhmästä. Arvioinnit koskivat potilaiden kokemia muutoksia mm. kyyvyssään tunnistaa masennusoireita, hallita kielteisiä ajattelutottumuksiaan ja toimia aiempaa rakentavammin lähisuhteissaan.

Tämä tutkimus on yksi monista, jotka ovat osoittaneet, miten pulmallisia psyykkisten ilmiöi-

Toisin kuin lääketutkimuksissa ns. vaikuttavaa ainesosaa ei psykoterapiassa voida erottaa kokeeseen osallistuvien ihmisten toiminnasta

den oiremittarit ovat käytännössä, mutta edelleen me uskomme jokseenkin sokeasti, että asteikon muotoon puserrettu itsearvio on jotenkin objektiivisempää tai tieteellisempää tietoa kuin henkilön omaa näkökulmaa valottava haastatteluaineisto.

Yhteistoiminta on vaikuttava ainesosa

Toisin kuin lääketutkimuksissa ns. vaikuttavaa ainesosaa ei psykoterapiassa voida erottaa kokeeseen osallistuvien ihmisten toiminnasta. Tämän on myös psykoterapian yhteisiin tekijöihin suuntautunut tutkimus osoittanut. Psykoterapia on kahden ihmisen yhteistoimintaa. Kumpikin vastaa toisen ilmaisuun, joten mikä tahansa tekninen toimenpide toteutuu kehittyvän vuorovaikutuksen osana. Tästä ns. responsiivisuuden ongelmasta johtuu myös se, että terapeutin toimenpiteitä ei ole mahdollista kuvata prosessimuuttujina, joiden avulla voitaisiin ennustaa terapian tuloksellisuutta. Terapeutin responsiivisuus tuhoaa tosiasiallisesti korrelaatioanalyysin tilastolliset edellytykset eli oletuksen terapeutti-, potilas- ja tulosmuuttujien riippumattomuudesta toisiinsa nähden. Tästä seuraa muun muassa se, että mikään prosessimuuttuja ei parhaimmillaankaan selitä edes viidennestä tuloksellisuusmuuttujien vaihtelusta.

Vaihtoehtoisia metodisia otteita

Hans Strupp tarkasteli edellä kuvattuja ongelmia jo arvioidessaan Vanderbilt I -tutkimusta 1980-luvun taitteessa, mutta psykoterapian tutkijoiden yhteisö ei vielä tuolloin ollut valmis näkemäänniiden merkitystä. Strupp julkaisi myös esimerkkejä siitä, miten yksittäisiä terapiaprosesseja tutkimalla voitiin osoittaa toimivien ja huonosti toimivien terapisuhteiden keskeisiä piirteitä. Hän muotoili tutkimustensa nojalla ns. PTO-periaatteen

(problem-treatment-outcome congruence), jota voidaan pitää kokeellisille ryhmävertailuille vaihtoehtoisena tutkimuksen metodologisena lähtökohtana.

PTO-periaatteen soveltaminen edellyttää, että terapeutin interventio perustuu ongelman potilaskohtaiseen muotoiluun. Sen mukaisesti hoidolle asetetaan ongelman kannalta relevantit tulostavoitteet ja määritään ne indikaattorit, joiden avulla muutosprosessi voidaan todentaa. Menettelytapa muistuttaa jo 1950-luvulla käyttäytymisterapioiden tutkimuksissa sovellettua ”single case”-metodiikkaa mutta ei ole käsitteellisesti sitoutunut ongelmien käyttäytymisteoreettiseen hahmotustapaan. PTO-periaatetta on sovellettu mm. psykodynaamisesti suuntautuneiden psykoterapioiden tutkimuksessa. Sen puitteissa on mahdollista soveltaa tutkimusmenetelmiä, joiden avulla potilaan ongelma voidaan muotoilla terapiamuodon kannalta relevanteilla tavoilla ja seurata siten terapiaprosessin sisältöjä ja tuloksellisuutta. Tunnetuimpia ovat CCRT (core conflictual relationship theme) ja SASB (structural assessment of social behavior). Kumpikin menetelmä on jo synnyttänyt varsin omaleimaisen tutkimusperinteen.

Toinen ryhmätutkimuksille vaihtoehtoinen psykoterapian tuloksellisuustutkimuksen metodologinen suuntaus on William B. Stilesin yhteistyössä Sheffieldin psykoterapiatutkimusryhmän kanssa 1980- ja 1990-luvun taitteessa kehittelemä assimilaatiomalli. Se jakaa psykoterapiaprosessin ja erityisesti siinä toteutuvat muuntumat kahdeksaan kehitysvaiheeseen. Mallia on sovellettu mm. kognitiivisen, eksploraatiivisen ja psykoanalyttisesti suuntautuvan terapian prosessin analyysissä. Psykoterapian tuloksellisuus voidaan osoittaa kuvaamalla, millainen muodonmuutos potilaan ongelmissa on tapahtunut ja mihin kehitysvaiheeseen niiden työstämisessä hän on hoidon lopussa päässyt.

Assimilaatiomalliin perustuva tutkimus hyödyntää yleisesti jo aiemmin koottuja vertailevia aineistoja, joissa hoidot on dokumentoitu äänitein tai videoinnein. Se, mikä kokeellisissa asetelmissä näyttäytyy vain tilastollisena vaihteluna, on mahdollista purkaa auki tapaus tapaukselta. Potilaan ja terapeutin yhteistoiminnan kehittämisen tutkiminen valottaa, millaisten prosessien

kautta hoitoryhmien sisäinen vaihtelu on syntynyt. Tutkimalla näin tarkemmin onnistuneiden ja vähemmän onnistuneiden hoitojen prosessia, voidaan jälkikäteen tuottaa uutta tietoa esimerkiksi terapeuttisen yhteistyösuhteen ja hoidon tekniesten elementtien vuorovaikutuksesta. Tutkijoille on tarjolla valtava määrä psykoterapioista huolellisesti tehtyjä dokumentteja, jotka vain odottavat tarkempaa analysointia.

FT Antero Kiiänmaa, professori Sirkka Keinänen-Kiukaanniemi, PsL Aarno Laitila, professorit Johannes Lehtonen ja Pekka Larivaara, dosentti Timo Muhonen, EL Juho Nummenmaa, dosentti Helena Pihko, FT, EL Raimo Puustinen ja käsikirjoituksen kolme tieteellistä arvioijaa ovat kommentillaan ja rohkaisullaan myötävaikuttaneet kirjoituksen lopulliseen asuun. Lämpimät kiitokseni heille.

Artikkeli on julkaistu aiemmin Duodecim-lehden numerossa 2004; 120: 2645–53.

Kirjallisuutta

Bergin A E, Garfield S L. Handbook of psychotherapy and behavior change. New York: John Wiley, 1971.

Blatt S J, Zuroff D C, Bondi C M, Sanislow III C A. Short- and long-term effects of medication and psychotherapy in the brief treatment of depression: Further analyses of data from the NIMH TDCRP. *Psychother Res* 2000; 10: 215–34.

Elkin I, Shea M T, Watkins J T, ym. National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Arch Gen Psychiatry* 1989; 46: 971–82.

Eysenck H J. The effects of psychotherapy: an evaluation. *J Consulting Psychol* 1952; 16: 319–24.

Henry W P. Defining the self in an interpersonal context. Teoksessa: Muran J C, toim. Self-relations in the psychotherapy process. Washington: American Psychological Association 2001, s. 267–84.

Klein D F. Control groups in pharmacotherapy and psychotherapy evaluations. *Prevention and Treatment* 1997;1. www.journals.apa.org/prevention/volume1/97_a1.html, 21.10.2003

Leiman P M. Kontrolloidut kokeet psykoterapian tuloksellisuustutkimuksessa. *Sosiaalilääket Aikak* 1990; 27: 119–26.

Luborsky L. A note on Eysenck's article "The effects of psychotherapy: an evaluation." *Br J Psychol* 1954; 45: 129–31.

1 Vaikuttavuustutkimuksen pulmallisuus psykoterapiassa LEIMAN

Luborsky L, Popp C, Luborsky E, Mark D. The core conflictual relationship theme. *Psychother Res* 1994; 4: 172–83.

Luborsky L, Diguier L, Seligman D A, ym. The researchers own therapy allegiances: a "wild card" in comparisons of treatment efficacy. *Clin Psychol Sci Pract* 1999; 6: 95–106

Messer SB. Empirically supported treatments: cautionary notes. *Medscape General Medicine* 2002; 4. www.medscape.com/viewarticle/445082, 3.11.2003.

Norcross J C, toim. Psychotherapy relationships that work. Therapist contributions and responsiveness to patients. New York: Oxford University Press, 2002.

Osatuke K, Glick M J, Stiles W B, Shapiro D A, Barkham M, Greenberg L S. Therapeutic Progress In CCT And CBT: Smooth Versus Sawtoothed. 2003; 34th Annual Meeting of the Society for Psychotherapy Research, Ulm, Germany.

Rosenzweig S. A transvaluation of psychotherapy: a reply to Hans Eysenck. *J Abnorm Soc Psychol* 1954; 49: 298–304.

Rudolf G, Grande T, Dilg R, ym. Structural changes in psychoanalytic therapies – the Heidelberg-Berlin study on long-term psychoanalytic therapy (PAL). Kirjassa: Leunzinger-Bohleber M, Target M, (toim.). Long-term psychoanalytic treatment. Perspectives for therapists and researchers. London: Whurr Publishers, 2004.

Stiles W B, Shapiro D A, Elliot R. Are all psychotherapies equivalent? *Am Psychologist* 1986; 41: 165–80.

Stiles W B, Elliott R, Llewelyn S P, ym. Assimilation of problematic experiences by clients in psychotherapy. *Psychotherapy* 1990; 27: 411–20.

Stiles W B, Morrison L A, Haw S K, Harper H, Shapiro D A, Firth-Cozens J. Longitudinal study of assimilation in exploratory psychotherapy. *Psychotherapy* 1991; 28: 195–206.

Stiles W B. Drugs, recipes, babies, bathwater, and psychotherapy process-outcome relations. *J Consult Clin Psychol* 1994; 62: 955–9.

Stiles W B. When more of a good thing is better: reply to Hayes et al. *J Consult Clin Psychol* 1996; 64: 915–8.

Strupp H H, Hadley S W. Specific versus nonspecific factors in psychotherapy: a controlled study of outcome. *Arch Gen Psychiatry* 1979; 36: 1125–36.

Strupp H H. Success and failure in time-limited psychotherapy: a systematic comparison of two cases. Comparison I. *Arch Gen Psychiatry* 1980; 37: 595–603.

Strupp H H, Schacht T E, Henry W P. Problem-treatment-outcome congruence: a principle whose time has come. Teoksessa: Dahl H, Kächele H, toim. Psychoanalytic process research strategies. New York: Springer, 1988, s. 1–14.

Strupp H H. The Vanderbilt I study revisited. *Psychother Res* 1998; 8: 17–29.

Westen D, Morrison K. A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: an empirical examination of the status of empirically supported therapies. *J Consult Clin Psychol* 2001; 69: 875–99.

Jälkisanat

Mikael Leiman

Artikkelin kirjoittamisen aikoihin näyttöön perustuvan lääketieteen käytännöt alkoivat voimistua maamme psykiatrisessa terveydenhuollossa. Näytönastekatsauksia ja Käypä hoito -suosituksia ohjasi myös mielenterveyden häiriöiden alalla vankka luottamus satunnaisesti kliinisten kokeiden kykyyn tuottaa luotettavinta tietoa psykoterapioiden vaikuttavuudesta. Psykoterapia-tutkimuksessa tämä näkemys kyseenalaistettiin ensimmäisen kerran 1960-luvun lopulla ja uudelleen 1980-luvun puolivälissä. Kritiikin ydin kohdistui virheelliseen analogiapäätelmään, joka tapahtuu, kun lääketutkimuksissa sovellettava metodologia ulotetaan yhteistoiminnalliseen, kehittyvään prosessiin. Artikkelissa tarkastellaan psykoterapiatutkijoiden kritiikin perusteita satunnaisesti kliinisten kokeiden aineistojen uudeleena-analysointien valossa.

Tieteellinen keskustelu ei ole kuitenkaan pysynyt horjuttamaan kliinisten kokeiden asemaa, joka on tullut keskeiseksi osaksi terveyspalvelujen kehittämisen strategiaa.

Artikkelissani *Vaikuttavuustutkimuksen pulmallisuus psykoterapiassa* kerrataan psykoterapia-tutkimuksen vuosikymmenien kuluessa syntynyttä kokemusta kliinisten kokeiden ongelmista. Siinä ei tarkastella kehitystä, joka on tapahtunut tilastollisen tutkimusmetodii-kan piirissä. Lähestymistavaltaan uudenlainen metodiikka tarjoutui vaihtoehdoksi jo 1960-

luvun kriittisessä keskustelussa. Sen käytännöllinen arvo näyttäytyi vasta 1980-luvun lopulla, kun terveyspalvelujen maksajat alkoivat kiinnostua psykoterapian yksilöllisestä tuloksellisuudesta. Maksajan kannalta on melko yhdentekevää, onko jokin psykoterapiamenetelmä todettu kliinisissä kokeissa yleisesti ottaen vaikuttavaksi, kun menetelmän keski-määräinen hoitovaste tietyn diagnoosin saaneiden potilaiden ryhmässä on 30–50 %. Sen sijaan maksajatahon kannalta on olennaista saada mahdollisimman perusteltu arvio sille, ketkä tietyn diagnoosin saaneista potilaista tulevat todennäköisimmin hyötymään psykoterapiastaan.

Psykoterapiatutkimuksen alalla on jo 1990-luvun alusta alkaen kehitelty odotettavissa olevan hoitotuloksen ennustemalleja. Ne nojautuvat suurien potilaskohorttien tilastolliseen analyysiin, jonka varassa on voitu kehittää tuloksellisuutta ennustavia potilaan profiloitimenetelmiä. 2000-luvun puolella tutkijoiden kiinnostus on enenevässä määrin alkanut kohdistua tuloksellisuutta ennustavaan terapeuttiseen vuorovaikutukseen. Samalla psykoterapeuttien yksilöllinen kyvykyys saada aikaan suotuisia tuloksia on tullut osaksi tutkimusta. Tämä lähestymistapa näyttää nyt olevan psykoterapiatutkimuksen voimakkaasti kasvava suuntaus vielä selvemmin kuin artikkelin kirjoittamisen aikoihin.