



**TURUN
YLIOPISTO**

Tekoälyavusteiset hoitomenetelmät ahdistuneisuuden hoidossa

Kandidaatintutkielma

Emilia Pulli ja Amanda Välitalo

Ohjaaja: Mika Koivisto

Turun yliopisto

Psykologian ja logopedian laitos

19.5.2025

Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -järjestelmällä.

TURUN YLIOPISTO

Psykologian ja logopedian laitos

PULLI, EMILIA; VÄLITALO, AMANDA: Tekoälyavusteiset hoitomenetelmät ahdistuneisuuden hoidossa

Kandidaatintutkielma, 40 s., 1 liites.

Psykologia

Toukokuu 2025

Nopeasti kehittyvät tekoälytyökalut tarjoavat uusia mahdollisuuksia mielenterveyden hoidossa. Ahdistusoireet ovat yleistyneet esimerkiksi Suomessa ja muissa EU-maissa COVID-19-pandemian jälkeen, mikä tekee tekoälyn käytöstä osana ahdistuneisuuden hoitoa ajankohtaisen tutkimusaiheen. Ahdistuneisuuden hoitoa voidaan tukea muun muassa chattiboteilla tai hyödyntämällä tekoälyä hallinnollisiin tehtäviin.

Tavoitteenamme oli tutkia tekoälyavusteisen hoidon tehokkuutta ahdistuneisuuden hoidossa tavanomaiseen hoitoon verrattuna. Aihetta ei ole aiemmin tutkittu systemaattisesti aktiivisilla kontrolli-interventioilla, ja osa aiemmasta tutkimuksesta oli keskittynyt vain tiettyyn tekoälyn osa-alueeseen.

Tutkimus toteutettiin systemaattisena kirjallisuuskatsauksena. Tietokantahaut tehtiin PubMed- sekä EBSCO:n APA PsycInfo-, CINAHL-, Medline- ja SocINDEX-tietokannoissa. Sisäänotto- ja poissulkukriteerien mukaisesti artikkeleissa tutkittiin tekoälyavusteista ahdistuneisuuden hoitoa siten, että koeryhmä saa mahdollisimman samanlaista hoitoa kuin kontrolliryhmä, mutta koeryhmän hoidossa on hyödynnetty tekoälyä. Katsaukseen sisällytettiin 12 vertaisarvioitua, tieteellistä artikkelia, joista tuloksia päädyttiin lopulta analysoidaan kymmenestä. Tutkimuksen lopullinen otoskoko oli 875 henkilöä, ja tutkimuksia oli kolmesta eri maasta.

Katsauksen tulosten perusteella tekoälyavusteinen hoito ei vaikuttaisi olevan tavanomaista hoitoa tehokkaampaa ahdistuneisuuden hoidossa. Havaitsemillamme tuloksilla voi kuitenkin olla kliinistä merkittävyyttä, sillä tekoälyn käytöllä osana hoitoa voidaan säästää resursseja ja lisätä hoidon saavutettavuutta.

Katsaus antaa entistä kattavamman kuvan tekoälyn hyödyntämisestä mielenterveyden interventioissa. Aihe on ajankohtainen, minkä takia tekoälyn kehittyessä nopeasti on oleellista huolehtia tekoälyavusteisen hoidon eettisyydestä. Jatkotutkimuksissa voitaisiin tutkia terapeuttista allianssia ja transferenssia sekä puhemallien hyödyntämistä tekoälyterapiassa.

Asiasanat: tekoäly, ahdistus, ahdistuneisuushäiriöt, ahdistuneisuuden hoito, systemaattinen katsaus

Sisällys

1. Johdanto	4
1.1 Ahdistuneisuus ja sen hoito.....	4
1.2 Tekoälyn käyttö mielenterveyden hoidossa	5
1.3 Aiempi tutkimus.....	6
1.4 Tutkimuskysymykset	7
2. Menetelmät.....	7
2.1. Aineiston hankinta	7
2.2. Aineiston luokittelu.....	9
3. Tulokset.....	9
3.1 Tutkimuksen aineisto	9
3.2 Ahdistuneisuuden arvioinnissa käytetyt mittarit.....	13
3.3 Tekoälyavusteiset hoitomenetelmät	15
3.4 Tulosten arviointi	21
4. Pohdinta	23
4.1 Tuloksiin vaikuttavat tekijät.....	23
4.1.1 Heikkolaatuiset tutkimukset.....	23
4.1.2 Muut tuloksiin vaikuttavat tekijät	24
4.2 Tutkimuksen vahvuudet ja rajoitukset	25
4.3 Tulosten suhteutuminen aikaisempaan tutkimukseen.....	26
4.4 Tutkimuksen merkitys.....	26
4.4.1 Eettiset kysymykset.....	27
4.4.2 Jatkotutkimus- ja sovellusehdotukset.....	28
4.5 Yhteenveto	29
Lähteet.....	30
Liitteet	39
Liite 1.	39
Liite 2.	40

1. Johdanto

Ahdistus on yleinen mielenterveyden oire, joka koskee monia ihmisiä. COVID-19-pandemian jälkeen masennus- ja ahdistusoireista kärsivien ihmisten osuus väestöstä on kasvanut merkittävästi niin Suomessa kuin muissa EU-maissa (OECD & European Observatory on Health Systems and Policies, 2023). Suomessa nuorilla, ja erityisesti nuorilla naisilla, on havaittu lisääntyntä ahdistuneisuutta. Ilmiötä havainnollistaa myös Terveyden ja hyvinvoinnin laitoksen toteuttama Kouluterveyskysely, jonka perusteella yläkoululaiset ja opiskelijat voivat huominkin kuin ennen (Nurmilaakso & Åström, 2024). Erityisesti tyttöjen kokemat ahdistus- ja masennusoireet ovat lisääntyneet huomattavasti – noin joka kolmas 8. tai 9. luokalla, lukiossa tai ammattikoulussa opiskeleva tyttö koki kohtalaista tai vaikeaa ahdistuneisuutta viimeisimmän Kouluterveyskyselyn mukaan (Terveyden ja hyvinvoinnin laitos [THL], 2023).

Tekoälyn käyttö on viime vuosien aikana tullut yhä suosittumaksi ja tekoälytyökalut ovat kehittyneet nopeasti, joten on hyvin ajankohtaista tutkia, onko tekoälyn avulla mahdollista toteuttaa uudenlaisia interventioita mielenterveyden hoidossa. Erään määritelmän mukaan tekoälyksi voidaan luokitella järjestelmät, jotka käyttäytyvät älykkäästi analysoimalla ympäristöään ja toimimalla jossain määrin itsenäisesti saavuttaakseen tietyt tavoitteet (Sheikh, ym., 2023a). Tekoälyä voidaan hyödyntää mielenterveyden hoidossa monilla eri tavoilla. Näin ollen selvitämme tässä systemaattisessa katsauksessa tekoälyavusteisten hoitojen käyttöä ahdistuneisuuden hoidossa.

1.1 Ahdistuneisuus ja sen hoito

Ahdistus tai ahdistuneisuus tunnetilana on pelon tai huolestuneisuuden tunne, joka johtuu osin tietoisesta tai tiedostamattomasta menneisyyteen, nykyhetkeen tai tulevaisuuteen liittyvästä huolesta, ja johon liittyy emotionaalisia ja fysiologisia reaktioita (Koponen ym., 2024; Rovasalo & Eerola, 2023). Ahdistuneisuushäiriöistä puhutaan silloin, kun ahdistuneisuus on voimakasta, pitkäkestoista ja normaalia arkielämää haittaavaa. ICD-10-tautiluokitus (World Health Organization [WHO], 2016) käsittää ahdistuneisuushäiriöiksi muun muassa yleistyneen ahdistuneisuushäiriön (F41.1), paniikkihäiriön (F41.0) ja sosiaalisten tilanteiden pelon (F40.1). Ahdistuneisuutta liittyy myös muihin psyykkisiin häiriöihin, kuten pakko-oireiseen häiriöön (F42), reaktioihin vaikeaan stressiin ja sopeutumishäiriöihin (F43) sekä masennustiloihin (F32). Monesti ahdistuneisuushäiriöihin liittyy niin ikään somaattista samanaikaissairastavuutta.

Käypä hoito -suosituksen mukaan ahdistuneisuushäiriöiden hoidossa voidaan käyttää psykososiaalisia hoitomuotoja, kuten psykoterapiaa ja lääkehoitoa, sekä niiden yhdistelmää (Koponen ym., 2024). Myös omahoito ja terveellisten elintapojen ylläpito sekä psykoedukaatio ovat osa ahdistuneisuuden hoitoa.

Yksi ahdistuksen hoidossa käytetyimmistä psykoterapioista on kognitiivinen käyttäytymisterapia (CBT, engl. cognitive behavioral therapy) (ks. esim. Beck, 2020). Tutkimuksien mukaan CBT on tehokasta esimerkiksi yleistyneen ahdistuneisuushäiriön hoidossa (Gale & Davidson, 2007) sekä paniikkikohtausten ja välttämiskäyttäytymisen vähentämisessä (Gould ym., 1995).

Perinteisesti CBT ja muut psykoterapiat toteutetaan yksilö- tai ryhmäkäynteinä kasvotusten psykoterapeutin kanssa. Tähän toteutusmuotoon liittyy kuitenkin haasteita, kuten kallis hinta, pitkät hoitojonot, sosiaalinen stigma ja pula pätevistä terapeuteista etenkin alueilla, joilla palveluiden saatavuus on heikko (Webb ym., 2017). Keinona näihin haasteisiin vastaamisessa voidaan käyttää teknologiaa, kuten internet-pohjaista (iCBT) tai tietokonepohjaista kognitiivista käyttäytymisterapiaa (CCBT) (Jiang ym., 2024). Esimerkiksi internet-välitteinen CBT voi olla paniikkihäiriön hoidossa yhtä tehokasta mutta kustannusvaikuttavampaa verrattuna kasvokkain toteutettuun CBT:hen (Bergström ym., 2010).

1.2 Tekoälyn käyttö mielenterveyden hoidossa

Aiempien psykoterapiatarkoitukseen hyödynnettyjen teknologioiden rinnalle on viime vuosien aikana tullut tekoäly. Tekoälyä voidaan hyödyntää monin eri tavoin mielenterveyden häiriöiden arvioinnissa, diagnosoinnissa ja hoidossa. Esimerkiksi suuriin kielimalleihin (LLM, engl. large language models) perustuvia, tekstiä tuottavia GPT-malleja (engl. generative pre-trained transformer) on alettu käyttää mielenterveyden häiriöiden hoidossa (Jiang ym., 2024). Spesifisti psykoterapiakäyttöön on kehitetty chattibotteja eli virtuaalisia keskusteluagentteja, jotka kykenevät jäljittelemään psykoterapeuttien sanallista ja sanatonta viestintää tarjoten tehokasta ja yksilöllistä hoitoa.

Aiemmat chattibotit ovat olleet sääntöpohjaisia, eli ne ovat toimineet jonkin ennalta rakennetun keskustelurungon mukaisesti. ELIZA (Weizenbaum, 1966) on eräs varhaisimmista sääntöpohjaisista chattiboteista, jonka voitaisiin katsoa sopivan terapiakäyttöön. ELIZA matki asiakaskeskeistä psykoterapeuttia etsimällä ihmisen kirjoittamasta tekstistä avainsanoja ja vastaamalla niiden avulla ihmiselle. Myöhemmin kehitetyt, tekoälyllä toimivat chattibotit käyttävät luonnollisen kielen prosessointia (NLP, engl. natural language processing) (Sheikh, ym., 2023b) aiempien sääntöpohjaisten järjestelmien ohella.

Tekoälyllä voidaan myös täydentää perinteisempiä hoitomuotoja esimerkiksi hyödyntämällä sitä datan analysoimisessa ja terapeuttien tukemisessa. Tekoälyä käyttämällä voidaan tämän lisäksi mahdollisesti vähentää hoitoon liittyviä vinoumia (Beg ym., 2024). Tällaiset tekoälytyökalut perustuvat usein koneoppimiseen (engl. machine learning), jolla tarkoitetaan tietokoneen kykyä oppia datasta laskennallisten algoritmien avulla itsenäisesti, ilman erillistä ohjelmointia (Helm ym., 2020).

Kun algoritmia on toistettu ja muokattu riittävästi, ohjelma pystyy syötteen (engl. input) perusteella ennustamaan tuloksen (engl. output) ja vertailemaan sitä ennalta tuntemiinsa tuloksiin.

Tekoälyn hyödyt ahdistuneisuuden hoidossa liittyvät etenkin sen saavutettavuuteen ja kustannustehokkuuteen. Digitaalisilla mielenterveyspalveluilla voidaan edistää saavutettavuutta, ja toisaalta tarpeiden mukaiset mielenterveyspalvelut ovat tärkeitä mielenterveyden häiriöihin liittyvän stigman vähentämisessä (Vorma ym., 2020). Tekoölyavusteisilla hoitomenetelmillä voidaan siis toteuttaa kustannustehokasta hoitoa, joka on helposti saavutettavissa ja sovellettavissa myös matalan kynnyksen palveluihin, ja niiden avulla voidaan täydentää perinteistä hoitoa.

1.3 Aiempi tutkimus

Tekoölyä käyttävistä hoitomuodoista on kattavasti tutkimusnäyttöä (ks. esim. Farzan ym., 2025; Zhong ym., 2024). Kuitenkin useassa tutkimuksessa tekoälyn vertailuinterventiona on käytetty passiivisia hoitomuotoja, jolloin erot hoidon tehokkuudessa voivat tekoälyn käytön sijaan johtua hoidon aktiivisuudesta. Tekoölyavusteisia hoitomuotoja on tutkittu erityisesti masennuksen ja myös ahdistuneisuuden hoidossa (ks. esim. Beg ym., 2024), mutta tarkempi tutkimusnäyttö etenkin ahdistuneisuuden tekoölyavusteisesta hoidosta olisi tarpeen. Monet aiemmat tutkimukset ja katsaukset keskittyvät niin ikään vain yhteen tekoälyn osa-alueeseen, kuten chattibotteihin tai koneoppimiseen.

Chattibotteja käsittelevässä 11 tutkimuksen systemaattisessa katsauksessa ja meta-analyysissä huomattiin, että chattibotti-interventiot olivat kontrolli-interventioita tehokkaampia lievittämään masennusoireita keskisuurella efektikoolla (Lim ym., 2022). Tässä tutkimuksessa kuitenkin keskityttiin pelkästään chattibotteihin, ja monissa mukana olleissa tutkimuksissa käytettiin ainoana kontrollimenetelmänä passiivisia kontrolli-interventioita.

Tutkimusaiheeseemme liittyen on tehty narratiivinen katsaus (Das & Gavade, 2024), jonka mukaan tekoölyavusteisen hoidon merkittävimmät hyödyt olivat hoidon parempi saatavuus, yksilöllisyys ja hallinnollisten toimien tehokkuus. Tekoölyä käytettäessä esiin nousee kuitenkin kysymyksiä, jotka liittyvät esimerkiksi yksityisyyden suojaan ja ihmisläheisen otteen säilyttämiseen hoidossa. Koska kyseinen kirjallisuuskatsaus oli narratiivinen, hoitomuotoja kuvailtiin lähinnä sanallisesti ja tutkimuksen tavoitteena oli kartoittaa tutkimusnäyttöä ja tekoölyyn liittyviä trendejä ilman tarkkoja tutkimuskysymyksiä. Siten aihetta olisi perusteltua tutkia vielä systemaattisen katsauksen menetelmin.

Aiemman tutkimuksen perusteella tekoölyavusteisilla hoitomuodoilla voi siis olla useita hyötyjä. Epäselvää on kuitenkin vielä, kuinka tehokkaita tekoölyinterventiot ovat verrattuna tyypillisiin hoitomuotoihin. Tutkimustietoa tekoölyavusteisista hoitomuodoista ahdistuneisuuden hoidossa on

vielä vähäisesti, ja tarvetta on etenkin tutkimukselle, jossa koe- ja kontrolli-interventiot ovat aktiivisia hoitomuotoja.

1.4 Tutkimuskysymykset

Aikaisemmissa tutkimuksissa on tutkittu tekoälyavusteisia hoitoja, mutta tällä hetkellä tutkimusnäyttöä ei juurikaan ole tekoälyavusteisten hoitojen tehokkuudesta, kun tekoälyä verrataan aktiiviseen interventioon. Aiemmat katsaukset ovat keskittyneet enemmän tiettyyn tekoälyn osa-alueeseen sen sijaan, että tarkasteltaisiin yleisesti tekoälyn käyttämisen vaikutusta. Myös hoitomuodon tehokkuutta spesifisti ahdistuneisuuden hoidossa on syytä tutkia enemmän. Tutkimme, miten tekoälyinterventiot vaikuttavat hoidon tehokkuuteen, ja toisaalta, mitä lisättävää tekoälyllä mahdollisesti on tavanomaisiin hoitomuotoihin.

Tutkimuksen tavoitteena on selvittää, miten tekoälyavusteiset psykoterapiamenetelmät ja muut hoitomuodot soveltuvat ahdistuneisuuden hoitoon. Tutkimuskysymyksemme on seuraavanlainen: Miten tekoälyn lisääminen ahdistuneisuuden hoitoon vaikuttaa hoidon tehokkuuteen?

2. Menetelmät

2.1. Aineiston hankinta

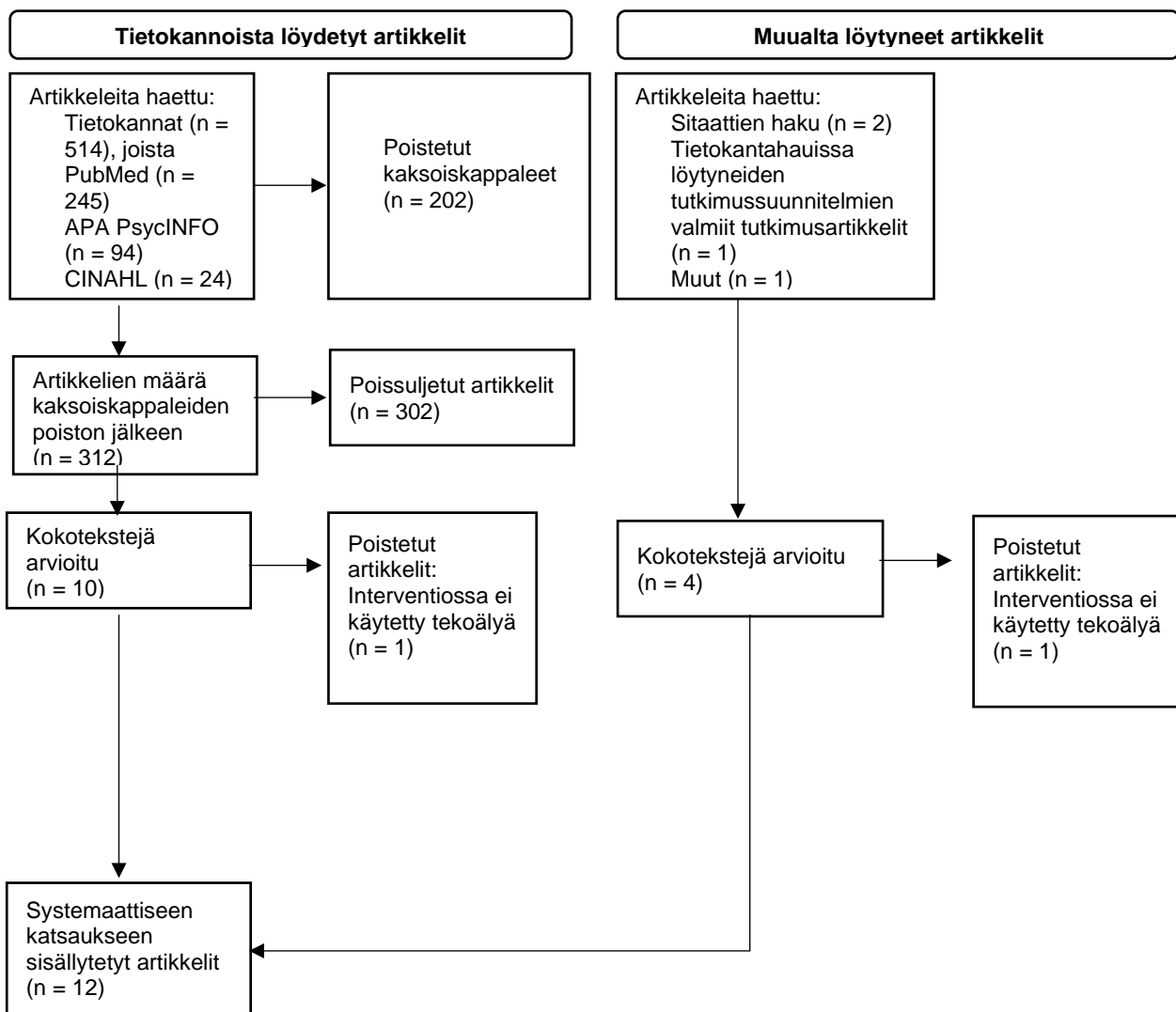
Tutkimus toteutettiin systemaattisena kirjallisuuskatsauksena. Tietokantahaut suoritettiin 11.2.2025 PubMed- sekä EBSCO:n APA PsycInfo-, CINAHL-, Medline- ja SocINDEX-tietokannoissa. EBSCO:n tietokannoissa hakiessa käytettiin vertaisarviointi-suodatinta, kun taas PubMedissä vastaavaa suodatinta ei ollut tarjolla. Hakulausekkeissa käytettiin tekoälyyn, psykoterapioihin ja ahdistuneisuuteen liittyviä hakusanoja. Tarkat tietokantakohtaiset hakulausekkeet ovat Liitteessä 1.

Sisäänottokriteerit olivat seuraavanlaiset: 1) artikkelin tulee olla vertaisarvioitu, tieteellinen artikkeli, 2) artikkeli on suomen- tai englanninkielinen, 3) artikkelissa tutkitaan tekoälyavusteisen hoitomuodon tehokkuutta ahdistuneisuuden hoidossa, 4) ahdistuneisuus on operationalisoitu ja 5) kontrolliryhmä saa mahdollisimman samankaltaista hoitoa kuin koeryhmä, erottavana tekijänä on tekoälyn käyttö. Poissulkukriteerit olivat puolestaan seuraavat: 1) artikkeli keskittyy johonkin muuta teknologiaa (ei tekoälyä) hyödyntävään hoitoon, 2) artikkelin tyyppi on katsausartikkeli tai tapaustutkimus ja 3) artikkelissa tarkastellaan hoidon vaikutusta pelkästään stressiin eikä varsinaisesti ahdistukseen.

Pyrimme ottamaan katsaukseen mukaan tutkimuksia, joiden kontrollitilanne oli niin sanottu aktiivinen kontrolli. Koska olemme kiinnostuneita nimenomaan siitä, miten tekoälyn lisääminen eri hoitomuotoihin vaikuttaa hoidon tehokkuuteen, on mielekästä vertailla keskenään sellaisia interventioita, joiden välillä on mahdollisimman vähän eroavaisuuksia tekoälyyn liittyvien seikkojen ulkopuolella. Passiivisilla kontrollitilanteilla, kuten psykoedukaatiolla tai odotuslistakontrollilla, on

mahdollista havaita suurempia eroavaisuuksia eri ryhmien välillä, mutta tällöin eroavaisuuksia voi selittää myös erot interventoiden aktiivisuudessa. Tämän takia systemaattiseen katsaukseen ei otettu mukaan tutkimuksia, joissa ainoa kontrollitilanne on passiivinen kontrolli. Passiivisen kontrollitilanteen hyväksyimme siinä poikkeustilanteessa, että myös koeryhmän saama interventio oli passiivinen, kuten omahoito, eli sisäänottokriteerien mukaisesti koe- ja kontrolliryhmät saivat mahdollisimman samankaltaista hoitoa.

Tietokantahakujen tuloksena saatiin yhteensä 514 artikkelia, joista kaksoiskappaleiden poiston jälkeen Rayyan-ohjelmassa (Ouzzani ym., 2016) jäi jäljelle 312 artikkelia. Artikkelien otsikot ja abstraktit käytiin läpi sokkoutetusti Rayyanissa sisäänotto- ja poissulkukriteerit huomioiden, minkä jälkeen kävimme yhdessä läpi artikkelit, joita emme olleet yksimielisesti sisällyttäneet tai poissulkeneet katsauksesta. Teimme keskustellen päätökset mukaan otettavista artikkeleista. Tämän jälkeen luimme valikoituneiden artikkeleiden kokotekstit ja karsimme pois yhden, koska sen interventiossa ei käytetty tekoälyä. Lisäksi saimme neljä artikkelia muita reittejä pitkin: yhden tutkimuksen havaitsimme katsausartikkelin lähdeluettelosta, ja toinen artikkeli meillä oli etukäteen tiedossa, mutta sitä ei jostain syystä löytynyt tekemillämme tietokantahauilla. Yksi artikkeli löytyi tietokantahausta saadun artikkelin lähdeluettelosta. Tietokantahauissamme saatiin useampi tutkimusprotokolla, joten selvitimme vielä, jos niistä löytyisi lopulliset tutkimukset. Tätä kautta havaitsimme vielä yhden artikkelin, joka oli julkaistu vasta tietokantahakumme jälkeen, sisällytettäväksi katsaukseen. Näistä neljästä muualta löytyneestä artikkelista yksi poissuljettiin tutkimuksesta. Lopulliseen katsaukseen valikoitui siis 12 artikkelia. Artikkelien valintaprosessi on havainnollistettuna Kuvassa 1.



Kuva 1. Vuokaavio artikkelien sisällyttämisestä katsaukseen. Vuokaavion tekemisessä on hyödynnetty Prisma 2020-mallikaavioita (Page ym., 2021).

2.2. Aineiston luokittelu

Seuraavaksi katsaukseen valikoidut artikkelit luokiteltiin olennaisten tietojen mukaan, ja tiedot kerättiin työskentelytaulukoihin. Näitä tietoja olivat julkaisutiedot, tutkimusasetelma, kuvaus otoksesta ja otoskoko, ahdistuneisuuden mittaamiseen käytetyt mittarit, toteutettu interventio ja kontrolli-interventio sekä tiivistelmä päätuloksista.

3. Tulokset

3.1 Tutkimuksen aineisto

Tutkimukseen sisällyttämämme aineisto on esitelty Taulukossa 1. Tutkimukset on julkaistu vuosina 2017–2025. Tutkittavia oli yhteensä 922 ja pudokkaita 201. Tutkittavien ikäjakauma vaihteli alakouluikäisistä ikääntyneisiin. Yleisin ikäryhmä oli työikäiset, joita tutkittiin yhdeksässä eri tutkimuksessa. Näistä kahdessa mielenkiinnon kohteena olivat yliopisto-opiskelijat. Lisäksi yhdessä tutkimuksessa tutkittiin kouluikäisiä, yhdessä alaikäisiä nuoria ja yhdessä ikääntyneitä.

Taulukko 1.

Tutkimuksen aineisto.

Julkaisutiedot	Otoskoko (pudokkaiden määrä)	Ikä vuosissa, M (SD)	Otoksen kuvaus
Beidel ym. (2021), Yhdysvallat	n = 38 (4) koeryhmä n = 25 (1) kontrolliryhmä n = 13 (3)	vaihteluväli 7–12 koeryhmä 9.35 (1.83) kontrolliryhmä 9.88 (1.86)	Kouluikäisiä lapsia, joilla ensisijaisena diagnoosina sosiaalisten tilanteiden pelko. 62 % naisia, 38 % miehiä.
Danieli ym. (2021), Italia	n = 21 koeryhmä n = 11 kontrolliryhmä n = 10	47.76 (8.07) vaihteluväli 33–61 koeryhmä 46.9 (5.89) kontrolliryhmä 48.7 (10.21)	Töissä käyviä aikuisia, joilla lievää-vaikeaa stressiä tai lievää-keskivaikeaa ahdistuneisuutta tai masennusta. 81 % naisia, 19 % miehiä.
Danieli ym. (2022), Italia	n = 45 (9) CBT ¹ + tekoäly n = 12 (4) CBT n = 15 (1) Pelkkä tekoäly n = 8 (2) Odotuslista-kontrolli n = 10 (2)	55.58 (5.08) CBT + tekoäly 55.17 (3.69) CBT 54.08 (4.11) Pelkkä tekoäly 55.63 (4.50) Odotuslista-kontrolli 57.20 (7.96)	Töissä käyviä aikuisia, joilla lievää-keskivaikeaa ahdistuneisuutta ja stressioireita. 78 % naisia, 22 % miehiä.
Gleason ym. (2025), Yhdysvallat	n = 141 (40) koeryhmä n = 49 (22) kontrolliryhmä n = 52 (18)	14.63 (1.21) vaihteluväli 13–17 koeryhmä 14.55 (1.19) kontrolliryhmä 14.71 (1.23)	Nuoria, jotka hoidettavana avohoidon klinikalla masennuksen ja/tai ahdistuneisuuden takia. 78 % naisia, 22 % miehiä.
Johnson ym. (2024), Australia	n = 82 (33) koeryhmä n = 43 (15) kontrolliryhmä n = 39 (18)	20.65 (2.38)	Psykologian opiskelijoita, jotka kokivat perfektionismin olevan ongelma elämässään. 84 % naisia, 12 % miehiä, 4 % ei-binäärisiä.
Kamita ym. (2019), Japani	n = 27 koeryhmä n = 15 kontrolliryhmä n = 12	koeryhmä 24.80 (1.57) kontrolliryhmä 25.33 (3.37)	Yliopisto-opiskelijoita, joista 56 % naisia, 44 % miehiä.
Leo ym. (2022), Yhdysvallat	n = 153 ryhmä A n = 51 ryhmä B n = 51 ryhmä C n = 51	55 (15) vaihteluväli 18–86	Aikuisia, jotka ortopedisessä hoidossa tuki- & liikuntaelimestön kivun vuoksi, ja joilla masennus- tai ahdistusoireita. 84 % naisia, 16 % miehiä.
Matsumoto ym. (2021), Japani	n = 20 (50) A (VR ² + chattibotti) = 6 (17) B (VR) = 7 (17) C (chattibotti) = 7 (16)	ei tiedossa	Vuorotyötä tekeviä sairaanhoitajia, joilla työn aiheuttamaa psykososiaalista kuormitusta, kuten tunnetyötä ja ylityötä.

Petersen ym. (2017), Yhdysvallat	n = 61 koeryhmä n = 35 kontrolliryhmä n = 26	83.4 koeryhmä 83.5(5.8) kontrolliryhmä 83.3(6.0)	Hoitokodissa asuvia ikääntyneitä, joilla lievä–keskivaikkea dementia. 77 % naisia, 23 % miehiä.
Russo ym. (2024), Italia	n = 160 koeryhmä n = 80 kontrolliryhmä n = 80	37.41 (6.77) vaihteluväli 25–49 koeryhmä 37.43 (6.74) kontrolliryhmä 37.40 (6.69)	PTSD ³ -kaltaisista oireista kärsiviä aikuisia, joilla on ollut pitkittynyt koronatauti. 50 % naisia, 50 % miehiä.
Sadeh-Sharvit ym. (2023), Yhdysvallat	potilaat n = 22 (25) koeryhmä n = 11 (12) kontrolliryhmä n = 11 (13) terapeutit koeryhmä n = 6, hoitivat 1–6 potilasta kontrolliryhmä n = 8, hoitivat 1–4 potilasta	potilaat 30.64 (11.02) koeryhmä 28.22 (9.46) kontrolliryhmä 32.96 (12.07)	Aikuisia, joilla ensisijaisena diagnoosina lievä tai keskivaikkea ahdistuneisuus tai masennus, ja jotka ohjattu avohoidon CBT:hen, sekä heitä hoitaneet terapeutit. Potilaista 72 % naisia, 28 % miehiä.
Suharwardy ym. (2023), Yhdysvallat	n = 152 (40) koeryhmä n = 68 (28) kontrolliryhmä n = 84 (12)	34	Vastikään synnyttäneitä naisia, jotka olivat keskimäärin oireettomia ahdistuneisuuden ja masennuksen suhteen.

¹CBT = kognitiivinen käyttäytymisterapia, ²VR = virtuaalitodellisuus, engl. virtual reality, ³PTSD = posttraumaattinen stressihäiriö.

Enemmistö katsauksemme tutkittavista oli naisia. Naisten osuus tutkimuksissa vaihteli välillä 50–100 %, ja yhdessä tutkimuksessa sukupuolijakauma ei ollut saatavilla.

Koska halusimme tutkia yleisesti ahdistusta, katsaukseen otettiin mukaan tutkimuksia, joissa ahdistusta oli esiintynyt monenlaisissa elämäntilanteissa ja erilaisilla diagnooseilla. Esittelemme seuraavaksi lyhyesti, minkälaisia kohderyhmiä eri tutkimuksissa on käsitelty.

Beidel kollegoineen (2021) käsitteli tutkimuksessaan sosiaalisten tilanteiden pelkoa, joka on ICD-10-tautiluokituksen mukainen ahdistuneisuushäiriö (WHO, 2016). Perfektionismin on havaittu olevan yhteydessä masennus- ja ahdistuneisuushäiriöihin, kuten sosiaalisten tilanteiden pelkoon ja paniikkihäiriöön, sekä pakko-oireiseen häiriöön (ks. esim. Antony ym., 1998). Katsauksessamme perfektionismia tutki Johnson kumppaneineen (2024).

Masennuksen ja ahdistuneisuuden on huomattu esiintyvän usein yhdessä kroonisen kivun kanssa, ja ne voivat vaikuttaa negatiivisesti ortopedisen hoidon vasteeseen esimerkiksi aiheuttamalla huonompaa fyysistä toimintakykyä toimenpiteen jälkeen, lisääntynyttä leikkauksen jälkeistä opioidien käyttöä ja heikentynyttä työhön paluun astetta. Aihetta tutkivat Leo ja kumppanit (2022).

Suharwardyn ja kollegoiden (2023) tutkimuksen osallistujiksi puolestaan valittiin vastikään synnyttäneitä naisia, sillä synnytys tuo mukanaan merkittäviä fyysisiä, psykologisia ja emotionaalisia muutoksia. Raskaudenaikainen ja synnytyksen jälkeen esiintyvä ahdistuneisuusoireilu on tavanomaista (Koponen ym., 2024).

Ahdistuneisuus on niin ikään yleinen oire dementiassa – arvioiden mukaan 8–71 %:lla dementoituneista esiintyy ahdistusoireita, ja 5–21 %:lla ahdistuneisuushäiriöitä (Kwak ym., 2017). Tyypillisiä ahdistusoireita dementiassa ovat esimerkiksi huolestuneisuus tulevaisuudesta, yliviireys, motorinen kiihtyneisyys ja nukahtamisvaikeudet. Dementiaan liittyen katsauksemme valikoitui Petersenin ja kollegoiden (2017) tutkimus.

Matsumoto ja kumppanit (2021) käsittelivät tutkimuksessaan sairaanhoitajien mielenterveyttä stressaavan vuorotyön yhteydessä. Kamita tutkimusryhmineen puolestaan (2019) tutkivat yliopistopöytäopiskelijoiden stressiä. Tutkimuksessa oli mitattu ahdistuneisuutta ja niin ikään muita oireita, kuten masentuneisuutta.

Mukaan valikoitui myös tutkimus, jossa tutkitaan tekoälyavusteisen EMDR (engl. eye movement desensitization and reprocessing, ”silmänliiketerapia”) -terapian tehokkuutta PTSD:n hoidossa (Russo ym. 2024). Vaikka PTSD:tä ei uusimman ICD-tautiluokituksen mukaan lasketa ahdistuneisuushäiriöksi (WHO, 2019), edellisessä ICD-10 luokituksessa PTSD ja erilaiset ahdistuneisuushäiriöt kuuluvat yhteiseen yläluokkaan ”Neuroottiset, stressiin liittyvät ja somatoformiset häiriöt” (WHO, 2016). EMDR-terapiaa voidaan käyttää myös

ahdistuneisuushäiriöiden hoidossa erityisesti, jos taustalla on traumaattisia muistoja (Haravuori, 2020). Nykyään on myös kehitetty uudenlaisia EMDR-tekniikoita, joissa pystytään käsittelemään ahdistuneisuushäiriöille tyypillisiä tulevaisuuden pelkoja (engl. flashforward images) (Thunnissen ym., 2024).

3.2 Ahdistuneisuuden arvioinnissa käytetyt mittarit

Seuraavaksi esittelemme tutkimuksissa käytetyt, ahdistuneisuuden arviointiin tarkoitetut mittarit. Mittarit on koottu taulukkoon 2.

Katsaukseen valikoituneissa tutkimuksissa GAD-7 (Generalized Anxiety Disorder Assessment) oli käytetyin mittari, sillä sitä käytettiin viidessä tutkimuksessa (42 % aineistosta). GAD-7 on seitsemän kysymyksen itsearviointimittari, joka noudattaa DSM-IV:n (American Psychiatric Association, 1994) kriteereitä yleistyneelle ahdistuneisuushäiriölle (Spitzer ym., 2006). GAD-7:ssä vastaaja arvioi ahdistusoireiden esiintymistä edeltäneiden 2 viikon aikana neliportaisella Likert-asteikolla. Pisteiden maksimimäärä on 21, jossa 0–4 kuvaa vähäistä ahdistuneisuutta, 5–9 lievää ahdistuneisuutta, 10–14 kohtalaista ahdistuneisuutta ja 15–21 vaikeaa ahdistuneisuutta.

SPAI-C (Social Phobia and Anxiety Inventory for Children) on 8–17-vuotiaille lapsille ja nuorille kehitetty sosiaalisten tilanteiden pelon vakavuuden ja laajuuden itsearviointimittari (Beidel ym., 1995). SPAI-C muodostuu 26 väittämästä, joita arvioidaan kolmiportaisella Likert-asteikolla. Vastaaja arvioi ahdistuneisuuttaan erilaisissa sosiaalisissa tilanteissa. SPAI-C:lla on erinomainen reliabiliteetti ja validiteetti sekä erottelevuus (Beidel ym., 2021). SPAIC-PV (Social Phobia and Anxiety Inventory for Children-Parent Version) (Higa ym., 2006) puolestaan on sosiaalisten tilanteiden pelosta kärsivien lasten tai nuorten vanhemmille tarkoitettu mittari. SPAIC-PV on niin ikään 26:n väittämän mittari kolmiportaisella asteikolla, ja sen kysymykset ovat samanlaiset kuin SPAI-C:ssä lukuun ottamatta sitä, että kysymyksen asettelu on vanhemman näkökulmasta. Higa ja kumppanit osoittivat, että SPAIC-PV korreloi merkitsevästi lapsen SPAI-C:ssä raportoiman sosiaalisen ahdistuksen kanssa ($r = .43, p < .01$).

CGI-asteikot (Clinical Global Impressions) (Guy, 1976) koostuvat kolmesta dimensiosta, joilla ammattilainen voi arvioida kliinistä yleisvaikutelmaa. Dimensiot ovat oireiden vaikeusaste, paraneminen ja hoidon tehokkuus (engl. severity of illness, global improvement and efficacy of treatment). Beidelin ja kumppaneiden (2021) tutkimuksessa CGI-asteikkoja käytettiin arvioimaan sosiaalisten tilanteiden pelon vaikeusastetta ja hoidon seurauksena tapahtuvaa paranemista.

SCL-90-R (Symptom Checklist-90-Revised) on moniulotteinen itsearviointikysely, jonka avulla voidaan arvioida stressiä ja erilaisten psykologisten oireiden voimakkuutta (Vaurio, 2011). Vastaaja arvioi asteikolla 0–4, kuinka paljon tietyt asiat ovat vaivanneet häntä kuluneen viikon aikana. Kyselyssä on 90 kysymystä, ja yhdeksän aladimensiota: somatisaatio, obsessiivis-kompulsiivisuus,

ihmissuhdeherkkyys, masentuneisuus, ahdistuneisuus, vihamielisyys, foobinen ahdistuneisuus, epäluuloiset ajatukset ja psykoottisuus (engl. somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism). Danielin ja kollegoiden tutkimuksessa (2021) ahdistuneisuuden mittaamiseen on käytetty nimenomaan ahdistuneisuuden aladimension tuloksia.

STAI (State-Trait Anxiety Inventory) on alun perin 1970-luvulla julkaistu ahdistuksen itsearviointimittari (Spielberger ym., 1970). 40:n väittämän mittari muodostuu tilanneahdistuksen (engl. state anxiety) ja piirreahdistuksen (engl. trait anxiety) neliportaisella Likert-asteikolla mitattavista osa-asteikoista, joissa kummassakin on 20 väittämää. Tilanneahdistus määritellään jännittyneisyyden, huolestuneisuuden sekä hermostuneisuuden tunteiden voimakkuutena tietynä ajankohtana, ja siihen liittyy autonomisen hermoston aktivaatio. Piirreahdistus puolestaan määritellään suhteellisen pysyvinä yksilöllisinä eroina ahdistusalttiudessa, mikä ilmenee sekä ahdistuneisuuden esiintymistiheytenä että todennäköisyytenä kokea tilanneahdistusta (Spielberger & Reheiser, 2009). STAI:n pistemäärä voi vaihdella välillä 20–80. Tässä katsauksessa STAI:ta hyödynnettiin Matsumoton (2021) sekä Kamitan ja kollegoiden (2019) tutkimuksissa, joissa pistemäärä 20–31 kuvasi vähäisempää, 32–34 kohtalaista, 35–41 korkeampaa ja 42–80 paljon korkeampaa ahdistuneisuutta.

PROMIS (Patient-Reported Outcomes Measurement Information System) -mittarit ovat PROMIS-terveysjärjestön kehittämiä itsearviointimenetelmiä (Yhdysvaltain terveysvirasto, 2023; THL, 2025). Leon ja kumppaneiden (2022) tutkimuksessa hyödynnettiin PROMIS-mittareista masentuneisuuden (v1.0), ahdistuneisuuden (v1.0), kivun häiritsevän vaikutuksen (v1.1) ja fyysisen toimintakyvyn (v2.0) mittareita. Näistä keskityimme katsauksessamme tarkastelemaan ahdistuneisuuden mittaria, jossa vastaaja arvioi ahdistuksen tunteitaan kuluneiden seitsemän päivän aikana viisiportaisella Likert-asteikolla.

RAID (Rating for Anxiety in Dementia) -mittari on tarkoitettu ahdistuneisuuden arviointiin dementian hoidossa (Shankar ym., 1999). Mittari on jaettu kuuteen aladimensioon; huolestuneisuus levottomuus ja valppaus, motorinen levottomuus, autonominen hyperaktiivisuus, fobiat ja paniikkikohtaukset (engl. worry, apprehension and vigilance, motor tension, autonomic hyperactivity, phobias, and panic attacks). Kyselyyn vastaa dementoituneen hoitaja tai omainen, joka arvioi ahdistusoireita kuluneiden kahden viikon aikana 5-portaisella Likert-asteikolla. Totaalipistemäärä voi vaihdella välillä 1–18, jossa yli 11 pisteen tulos kuvaa kliinisesti merkittävää ahdistuneisuutta.

IES-R (The Impact of Event Scale -Revised) (Horowitz ym., 1979; Weiss & Marmar, 1997) -mittarilla kartoitetaan traumaattiseen tapahtumaan liittyviä oireita. Kyselyssä on kolme dimensiota:

tunkeilevat ajatukset, välttely ja yliaktivoituminen (engl. intrusion, avoidance, and hyperarousal). Vastaaja arvioi asteikolla 0–4 22 kysymykseen vastaten, kuinka paljon tietyt oireet ovat vaivanneet häntä kuluneen viikon aikana.

3.3 Tekoälyavusteiset hoitomenetelmät

Tekoälypohjaisten teknologioiden laaja valikoima näkyi myös katsauksemme aineistossa, sillä tekoälyä oli hyödynnetty eri interventioissa hyvin eri tavoin. Hoitomenetelmät on niin ikään koottu Taulukkoon 2.

Käytetyin tekoälyn muoto oli chattibotit, joita hyödynnettiin kahdeksassa tutkimuksessa. Näissä tutkimuksissa kaikki chattiboteista toimivat ikään kuin terapeutin roolissa lukuun ottamatta Johnsonin ja kumppaneiden tutkimusta (2024), jossa tekoälyltä pystyi kysymään neuvoa ja tarkentavia kysymyksiä intervention sisällöstä ja vaiheista. Kahdessa tutkimuksessa hyödynnettiin tekoälyä, joka teki yhteenvetoja terapiakäynneistä ja keräsi dataa. Mukana oli myös tutkimus, jossa virtuaalitodellisuudessa esiintyvät hahmot toimivat tekoälyllä. Lisäksi yhdessä tutkimuksessa käytettiin robottilemmikkiä, jolle luotiin ”persoonallisuus” tekoälyn avulla. Seuraavaksi esittelemme tutkimuksissa käytetyt hoitomenetelmät ja niissä hyödynnetyt tekoälyteknologiat.

Kognitiivinen käyttäytymisterapia (CBT) on tutkituin ja eniten empiiristä tutkimusnäyttöä saanut psykoterapiamenetelmä, jota voidaan käyttää muun muassa masennuksen, yleisen ahdistuneisuuden, paniikkihäiriön, pakko-oireisen häiriön ja traumaperäisen stressihäiriön hoidossa (Fenn & Byrne, 2013). CBT:ssä käytetyt menetelmät voidaan jakaa kognitiivisiin ja käyttäytymismenetelmiin. Kognitiivisilla menetelmillä voidaan pyrkiä vaikuttamaan esimerkiksi maladaptiivisiin ydinuskomuksiin korvaamalla niitä uusilla, adaptiivisemmilla uskomuksilla. Käyttäytymismenetelmiä hyödynnetään etenkin ahdistuneisuushäiriöiden hoidossa, esimerkiksi katastrofijattelun tai välttelykäyttäytymisen hoidossa. CBT:n suosio psykoterapiamenetelmänä lienee myös syy siihen, että useassa katsauksemme tutkimuksessa käytetyksi menetelmäksi oli valittu nimenomaan CBT tai siihen pohjautuva hoito.

Danielin ja tutkimusryhmän (2021, 2022) kehittämä TEO (Therapy Empowerment Opportunity) on henkilökohtainen tekoälykeskusteluagentti (engl. mobile personal health care agent, mPHA). TEO perustuu stressinhallintaharjoitteluun (engl. stress management training, SMT) ja CBT:hen, ja se on suunniteltu stressin ja ahdistuneisuuden preventioon ja hoitoon. Danielin ja tutkimusryhmän aikaisemmassa tutkimuksessa TEOa käytettiin SMT-CBT:n tukena, ja Danielin ja tutkimusryhmän myöhemmässä tutkimuksessa yksi ryhmistä sai hoitoa pelkästään TEO:n avulla. Keskusteluagentti käyttää esimerkiksi CBT:ssä hyödynnettyä ABC-tekniikkaa (engl. antecedents, beliefs, consequences; *ennakkotapahtumat, uskomukset, seuraukset*) (Ellis, 1962), jossa pyritään ymmärtämään tapahtumien, ajatusten sekä tunteiden ja käyttäytymisen välistä yhteyttä. TEO-agentti

pystyy tunnistamaan tekstistä käyttäjänsä tunnetiloja, uskomuksia ja elämäntapahtumia, ja toimimaan luonnollisen kielen välityksellä vuorovaikutuksessa käyttäjän kanssa. Lisäksi TEO antaa käyttäjille henkilökohtaisia vinkkejä ja harjoituksia terapian edistymiseen perustuen, ja välittää terapeutille raportin sen kanssa käydyistä keskusteluista, jotta niitä voi hyödyntää terapiaprosessissa.

Woebot on yhdysvaltalaisen Woebot Healthin tekoälykäyttöinen mielenterveyschattibotti. Woebotin sisällöt perustuvat CBT:hen, dialektiseen käyttäytymisterapiaan ja intersoonalliseen psykoterapiaan (Woebot Health, 2025). Suharwardyn ja kollegoiden (2023) tutkimuksessa käytettiin Woebotia, joka kyseisessä tutkimuksessa keskittyy raskaudenjälkeisiin teemoihin. W-GenZD Gleasonin ja kollegoiden (2025) tutkimuksessa on niin ikään Woebotin chattibottiin pohjautuva relationaalinen agentti (engl. relational agent), joka tarjoaa tunteidenhallintatekniikoita, mielialan seuranta, empatiaa ja psykoedukaatiota.

Wysa-sovellus perustuu CBT:hen, mindfulnessiin ja motivoivaan haastatteluun (Wysa, 2024). Sovelluksessa on tekoälypohjainen chattibotti sekä mahdollisuus viestitellä ammattilaisen kanssa. Leon ja kumppaneiden (2022) tutkimuksessa Wysasta käytettiin versiota, jossa oli erityisesti kroonisesta kivusta kärsiville räätälöityjä lisäominaisuuksia.

Matsumoton (2021) ja Kamitan (2019) tutkimusryhmien interventiona oli digitaalinen SAT (engl. structured association technique). SAT-interventiossa käytetään useita eri tekniikoita, kuten hoidettavalle näytettävät erilaiset kuvat maisemista ja hymyilevistä kasvoista, joiden avulla vähennetään epämiellyttäviä tunteita ja luodaan turvallisuuden tunnetta. Menetelmän kehittänyt Munakata kuvailee artikkelissaan (2009) SAT:n olevan ”uuden sukupolven CBT”. Tekoälynä sekä Matsumoton että Kamitan tutkimusryhmien tutkimuksissa oli Japanin suosituimman sosiaalisen median palvelun, LINE:n, kautta toimiva chattibotti. Matsumoton tutkimusryhmän interventiossa oli tämän lisäksi VR-sovellus, joka ei suoranaisesti hyödynnä tekoälyä.

Johnsonin ja kumppaneiden (2024) käyttämä interventio oli omahoito (engl. self-help). Omahoitoon kuului 11-moduulinen työkirja *Changing Perfectionism (Perfektionismin muuttaminen)*. Koeryhmä sai ohjeet tekoälyn hyödyntämiseen omahoidon tukena. He saivat esimerkiksi kysymyksiä, joita kysyä tekoälyltä ja he saivat käyttää mitä tahansa tekoälytyökalua; tekoälyn käyttöön tottumattomille suositeltiin ChatGPT:tä. Kontrolliryhmä sai käyttöönsä vastaavan työkirjan, josta oli poistettu tekoälyn käyttöön liittyvät ohjeet, ja heitä pyydettiin nimenomaisesti olemaan käyttämättä tekoälyä.

SET-C (Social Effectiveness Therapy for Children) -terapiaan (Beidel ym., 2004) sisältyy ryhmämuotoista sosiaalisten taitojen harjoittelua, ryhmäaktiiviteetteja, tosielämässä tapahtuvaa yksilöaltistusterapiaa ja kotitehtäviä. Sosiaalisten taitojen ryhmämuotoinen harjoittelu koostui Beidelin ja tutkimusryhmän (2021) tutkimuksessa keskustelutaitojen, empaattisen kuuntelemisen taitojen, ystävyys- ja konfliktinratkaisutaitojen sekä assertiivisuuden harjoittelusta. Ryhmäaktiiviteetit

seurasivat aina sosiaalisten taitojen harjoittelua. Niissä sosiaalisten tilanteiden pelosta kärsivät lapset osallistuivat aktiviteetteihin, kuten keilailuun tai minigolfiin, samanikäisten, ei-ahdistuneiden ikätovereiden kanssa. Yksilöaltistusterapiassa puolestaan käsiteltiin kunkin osallistujan yksilöllistä pelkoa, kuten ääneen lukemista, soittimen soittamista tai puheen pitämistä. Koetilanteessa hyödynnettiin tekoälypohjaista Pegasys-VRTM-sovellusta (Wong Sarver ym., 2014), jossa oli edellä kuvattu SET-C:n rakenne, mutta ryhmäaktiviteetit ja kotitehtävät toteutettiin virtuaalitodellisuudessa. Sovelluksessa on virtuaalinen kouluympäristö ja 14 tekoälyllä toimivaa hahmoa, joiden kanssa käyttäjä on vuorovaikutuksessa.

Russo tutkimusryhmineen (2024) tutkivat etäyhteydellä toteutettavaa EMDR-terapiaa. Kontrolliryhmän EMDR-terapia toteutettiin kasvotusten. Koetilanteessa terapian tukena hyödynnettiin koneoppimiseen perustuvaa tekoälyalustaa, joka tallentaa ja analysoi potilaan silmänliikkeitä bilateraalissa stimulaatiossa. Interventioon kuului kolme 45 min tapaamista niin, että tapaamisten välissä oli korkeintaan viikko aikaa.

Sadeh-Sharvit ja kollegat (2023) käyttivät tutkimuksessaan Eleos-tekoälyalustaa, joka tukee kliinistä päätöksentekoa ja automatisoi hallinnollisia tehtäviä. Tekoäly tekee yhteenvetoja ja erilaisia analyyseja terapiaistunnoista, antaa terapeuteille palautetta ja tekee graafisia tiivistelmiä potilaan täyttämistä kyselyistä.

Petersenin ja kollegoiden (2017) käyttämä PARO-robottilemmikki on kehitetty vaihtoehtona eläinavusteiselle terapialle. PARO on hylkeenpoikasen näköinen robotti, joka jäljittelee eläimen käytöstä sekä kehittää oman ”luonteensa” ajan myötä tekoälyn avulla.

Taulukko 2.

Interventiot ja tulokset.

Tutkimus ja tutkimusasetelma	Ahdistuneisuuden mittarit	Interventio		Seuranta-ajankohdat	Tiivistelmä päätuloksista	Efektikoko
		Toteutettu interventio	Vertailumenetelmä			
Beidel ym. (2021) RCT ¹	SPAI-C ² , SPAIC-PV ³ , CGI ⁴	SET-C ⁵ , jonka apuna Pegasys-VR TM -sovellus. 12x 2 h sosiaalisten taitojen harjoittelua ja VR-ryhmäaktiiviteetteja, + 12x 1 h yksilöaltistusterapiaa, yht. 36 h, tapaamisia 2x/vko.	SET-C. 12x 3 h sosiaalisten taitojen harjoittelua ja ryhmäaktiiviteetteja (ylimääräinen tunti vrt. koeryhmään on paikasta toiseen liikkumista) + 12x 1 h yksilöaltistusterapiaa, yht. 48 h, tapaamisia 2x/vko.	Hoidon jälkeen	Oireet helpottuivat molemmissa ryhmissä, eikä ryhmien välillä ollut juurikaan eroja. Ainoa tilastollisesti merkitsevä ero ryhmien välillä oli, että SPAIC-PV-tulosten perusteella Pegasys-VR interventio oli tehokkaampi kuin kontrollitilanne.	M (SE) pre - post: SPAI-C Pegasys-VR TM 30.41 (2.01) - 18.18 (2.10), SET-C 28.56 (2.53) - 22.23 (2.63) SPAIC-PV Pegasys-VR TM 36.64 (1.73) - 21.34 (2.39), SET-C 32.23 (2.25) - 25.19 (3.11) CGI Pegasys-VR TM 2.89 (0.38) -4.71 (0.18), SET-C 2.46 (0.48) -4.45 (0.22)
Danieli ym. (2021) arviointitutkimus	SCL-90-R ⁶ : ahdistuneisuuden alasteikko	SMT-CBT ⁷ , jonka tukena keskusteluagentti, 8 tapaamista.	SMT-CBT etänä, 8 tapaamista.	8 vk, 3 kk	Ahdistusoireet helpottuivat tehokkaammin kontrolliryhmässä kuin koeryhmässä. Koeryhmässä oireet olivat pahentuneet alkumittauksen ja hoidon lopun välillä.	M(SD) pre – post. parametriset testit: SMT-CBT + keskusteluagentti 57.55 (10.21) - 59.27 (9.09) SMT-CBT 56.50 (9.17) - 49.60 (6.15)
Danieli ym. (2022) RCT	GAD-7 ⁸ ja SCL-90-R: ahdistuneisuuden alasteikko	8 vk CBT etänä, tukena TEO-keskusteluagentti.	8 vk CBT etänä, myös ryhmä, joka käytti pelkästään TEOa sekä odotuslistaryhmä.	4 vk, hoidon jälkeen (8 vk), 3 kk	Kaikissa ryhmissä tapahtui oireiden pahenemista 4 vk ja 8 vk mittausten välillä, jolloin koronatapausten määrä Italiassa kasvoi huomattavasti.	M(SD) pre - post: GAD-7 CBT+TEO 7.10 (3.64) - 4.50 (2.76) CBT 9.50 (5.23) - 7.08 (5.14) TEO 6.00 (4.43) - 5.00 (4.86) odotuslistaryhmä 3.14 (2.04) - 5.14 (1.77)
Gleason ym. (2025) RCT	GAD-7	W-GenZD-relaationaalinen agentti 4 vk ajan.	Ryhmämuotoinen CBT etänä, 4x 60 min yht. 4 vk ajan.	Hoidon jälkeen (4 vk), 8 vk	W-GenZD oli yhdenveroinen CBT:n kanssa masennusoireiden vähentämisessä, mutta	M (SD) pre - post: W-GenZD 10.84 (6.03) – 8.81 (5.49) CBT

						ahdistuneisuus pysyi samalla tasolla.	11.25 (5.30) – 9.79 (5.59)
Johnson ym. (2024) pilotti-RCT	GAD-7	Ohjattu omahoito: työkirja + ohjeet tekoälyn hyödyntämiseen, omaan tahtiin suoritettavaksi 4 vk aikana.	Pelkkä omahoito: työkirja omaan tahtiin suoritettavaksi 4 vk aikana.	Hoidon jälkeen (4 vk), 8 vk		Tekoälyn käyttö ei vaikuttanut intervention tehokkuuteen.	Ryhmien välinen ero: Cohenin d (95 % CI) = 0.14 (–0.38, 0.66)
Kamita ym. (2019) vertailututkimus	STAI ⁹	Chattibotti LINE-sovelluksessa. Digital-SAT ¹⁰ .	Web Digital-SAT.	Hoidon jälkeen (koko intervention kesto n. 1,5 h)		Tekoälyn käytöllä ei ollut vaikutusta ahdistuksen hoidossa.	M(SD) pre - post: CB 45.33 (8.14) - 40.80 (8.28), $p = 0.038$ Web 45.83 (12.68) - 40.33 (12.81), $p = 0.245$ Ryhmien välinen ero $p = 0.480$
Leo ym. (2022) retrospektiivinen kohorttitutkimus	PROMIS ¹¹ : ahdistuneisuus v1.0	A: Pääsy Wysa-mobiilisovellukseen 2 kk ajaksi. Lisäksi tavanomainen ortopedinen hoito.	B: Tavanomainen ortopedinen hoito ilman erityistä mielenterveysinterventiota. C: Tavanomainen ortopedinen hoito ja psykologin tapaamiset kasvokkain.	2 kk		Virtuaalisen hoidon tehokkuus oli verrattavissa kontrolliryhmään. Ryhmien välillä ei havaittu tilastollisesti merkitseviä eroja ahdistuksen suhteen.	MD (95 % CI) ryhmä C vrt. ryhmä A –1.9 (–4.5, 0.8), $p = .18$
Matsumoto ym. (2021) interventiotutkimus	STAI	A: Chattibotti LINE-sovelluksessa sekä VR-sovellus. VR yövuoron jälkeisinä päivinä ja chattibotti päivävuoropäivinä 4 vk ajan. Digital-SAT.	B: Pelkkä VR-sovellus yövuoron jälkeisinä päivinä 4 vk ajan C: Pelkkä chattibottisovellus yövuoron jälkeisinä päivinä ja päivävuoropäivinä 4 vk ajan. Digital-SAT.	Hoidon jälkeen (4 vk)		STAI ei eronnut missään ryhmässä alku- ja loppumittausten välillä. Ryhmiä ei verrattu tilastotestein toisiinsa. Lopullinen n hyvin pieni.	Mdn [25%, 75%] pre - post: CB+VR 50.0 [47.5, 53.5] - 51.0 [48.0, 59.3], $p = 0.345$, $z = -0.943$ VR 50.0 [42.0, 54.5] - 53.5 [38.8, 58.5], $p = 0.752$, $z = -0.316$ CB 55.0 [38.0, 59.0] – 53.0 [36.0, 59.0], $p = 0.752$, $z = -0.316$
Petersen ym. (2017) satunnaistettu lohkokoe	RAID ¹²	PARO-robotilemmikki kuuden henkilön ryhmässä, 20 min istunto 3x viikossa 3 kk ajan.	Hoitokotien tyypilliset aktiviteetit, kuten musiikki, liikunta ja erilaiset virikkeet, 20 min istunnoissa päivittäin 3 kk ajan.	Hoidon jälkeen (3 kk)		Ahdistuneisuus helpottui robotilemmikkiä käyttäneillä enemmän kuin kontrolliryhmällä. Ero oli tilastollisesti merkitsevä. Huomattiin myös, että robottia käyttäneiden kipu lievittyi ja tarve kipulääkitykselle väheni.	MD (SE) post - pre: PARO 2.5 (0.6) Vertailuryhmä 0.55 (0.2)
Russo ym. (2024) RCT	IES-R ¹³	EMDR ¹⁴ -terapia etänä ja silmänliikedataa keräävä tekoäly. 3x 45 min tapaamista niin, että tapaamisten välissä korkeintaan viikko aikaa.	EMDR-terapia kasvotusten. 3x 45 min tapaamista niin, että tapaamisten välissä korkeintaan viikko aikaa.	Hoidon jälkeen		Ei merkitseviä eroja tekoälyterapian ja normaalin terapian välillä hoidon tehokkuudessa.	Ryhmien välinen ero: Cohenin $d = 0.04$, $F = 0.09$, $p = 0.76$
Sadeh-Sharvit ym. (2023) RCT	GAD-7	CBT:n ohella hyödynnettävä Eleos-tekoälyalusta. Terapian ensimmäiset 2 kk.	CBT kasvokkain.	1 kk ja 2 kk		Ahdistusoireet vähenivät kummassakin ryhmässä, mutta oireet vähenivät tehokkaammin	Cohenin d (95 % CI) pre - post: Eleos 0.78 (-0.11, 1.62)

Suharwardy ym. (2023) ei-sokkoutettu RCT	GAD-7	Woebot-chattibotti sekä tavanomainen raskaudenjälkeinen hoito. Botin vapaa käyttö, kannustettiin käyttämään päivittäin.	Tavanomainen raskaudenjälkeinen hoito: seurantakerta 6 vk jälkeen ja mahdollinen mielenterveyden hoito tarvittaessa.	Synnytyksen jälkeen sairaalassa sekä 2, 4 ja 6 vk synnytyksen jälkeen	terapeutin käytettyä tekoälyalustaa.	CBT 0.14 (-0.70, 0.97)
					Ahdistuksessa ryhmien välillä ei ollut tilastollisesti merkitseviä eroja. Toisaalta kliinisen ahdistuksen raja ei ylittynyt koko otoksessa alkumittauksessa eikä 6 vk kohdalla.	MD (SD) pre - post: Woebot -0.72 (3.35) Vertailuryhmä .036 (3.78), $p = .19$

¹RCT = satunnaistettu kontrolloitu tutkimus, engl. randomized controlled trial, ²SPAI-C = Social Phobia and Anxiety Inventory for Children, ³SPAIC-PV = Social Phobia and Anxiety Inventory for Children-Parent Version, ⁴CGI = Clinical Global Impressions Scale, ⁵SET-C = Social Effectiveness Therapy for Children, ⁶SCL-90-R = Symptom Checklist-90-Revised, ⁷SMT-CBT = stress management training CBT, ⁸GAD-7 = Generalized Anxiety Disorder Assessment, ⁹STAI = State-Trait Anxiety Inventory, ¹⁰SAT = Structured Association Technique, ¹¹PROMIS = Patient-Reported Outcomes Measurement Information System, ¹²RAID = Rating for Anxiety in Dementia, ¹³IES-R = Impact of Event Scale-Revised, ¹⁴EMDR = silmänliikkeillä poisherkeittäminen ja uudelleenprosessointi, "silmänliiketerapia".

3.4 Tulosten arviointi

Katsaukseen mukaan otettujen tutkimusten tulosten mukaan tekoälyavusteisuus vähensi ahdistuneisuutta tai koe- ja kontrolliryhmien välillä ei havaittu eroja. Yhdessä tutkimuksessa (Danieli ym., 2021) tekoälyryhmän ahdistuneisuus lisääntyi hoidon alun ja lopun välillä, mutta tämä tulos selittyy mahdollisesti ulkoisilla tekijöillä.

Kahdessa tutkimuksessa (Kamita ym., 2019; Matsumoto ym., 2021) havaitsimme huomattavia laadullisia puutteita, joten päätimme jättää näiden tutkimusten tulokset huomiotta katsauksen kokonaistuloksissa. Kuvailimme tutkimuksia tarkemmin alaluvussa 4.1.1.

Tekoäly-intervention ja tavanomaisen hoidon välillä ei havaittu merkitseviä eroja seitsemässä tutkimuksessa (Beidel ym., 2021; Danieli ym., 2022; Gleason ym., 2025; Johnson ym. 2024; Leo ym., 2022; Russo ym., 2024; Suharwardy ym., 2023). Edellä mainituissa tutkimuksista kuudessa tapahtui ahdistusoireiden vähenemistä alku- ja loppumittauksen välillä. Danielin ja kollegoiden tutkimuksessa taas tapahtui oireiden pahenemista molemmissa ryhmissä hoidon puolenvälin ja loppumittauksen välillä. Johnsonin ja kumppaneiden sekä Russon ja kumppaneiden tutkimuksissa raportoitu efektikoko ei ylittänyt pienen efektin rajaa.

Tekoälyavusteisuus vähensi ahdistuneisuutta tilastollisesti merkitsevästi verrattuna kontrolliryhmään kahdessa tutkimuksessa (Petersen ym., 2017; Sadeh-Sharvit ym., 2023). Näissä tutkimuksissa käytetyt tekoälysovellukset olivat melko erilaisia; Sadeh-Sharvitin ja kollegoiden tutkimuksessa tekoäly hoiti hallinnollisia tehtäviä ja teki yhteenvetoja, kun taas Petersenin tutkimusryhmä tutki robottilemmikkiä.

Osa tutkimuksista, joissa tekoäly- ja kontrolliryhmien välillä ei ollut eroa, oli esitelty ns. ”non-inferiority” -tutkimuksina (Beidel ym., 2021; Gleason ym., 2025), jolloin tutkimuksen tavoitteena on selvittää, että koeinterventio ei ole kontrolli-interventiota huonompi. Tällöin tekoälyavusteisella terapiamuodolla voi olla kliinistä merkittävyyttä, sillä tekoälyä hyödyntävän terapian toteuttaminen voi säästää resursseja, ja sen avulla voidaan mahdollisesti lisätä hoitomuotojen saavutettavuutta henkilöille, jotka eivät jostain syystä pääse hoitoon paikan päälle. Tekoällyn käyttö mahdollistaa niin ikään esimerkiksi hallinnollisten tehtävien ulkoistamisen. Toisaalta Beidelin ja kollegoiden tutkimuksen interventiossa oli osia, joissa tekoälyä ei käytetty lainkaan, joten voi myös olla mahdollista, että kyseinen intervention osa on muutenkin tehoton eikä tekoällyn lisääminen silloin tuota lisähyötyjä.

Danieli ja kumppanit olettivat tutkimuksissaan (2021, 2022) havaitun oireiden pahenemisen johtuneen Italian koronapandemiatilanteen vaikeudesta. Aiemmassa tutkimuksessa saadun tuloksen mukaan ahdistuneisuus väheni tehokkaammin kontrolliryhmässä, joka sai pelkkää tavanomaista psykoterapiaa verrattuna ryhmään, joka käytti tekoälyä psykoterapian tukena. Intervention jälkeen

koeryhmän ahdistuneisuus SCL-90-R:llä mitattuna oli koeryhmällä korkeammalla tasolla kuin kontrolliryhmällä. Tutkimusryhmä oletti intervention lopun ajoittumisen koronarajoitusten alkuun vaikuttaneen todennäköisesti tutkimustuloksiin. Toisaalta tämä ei selitä sitä, miksi koeryhmässä ahdistuneisuus oli lisääntynyt enemmän, sillä koe- ja kontrolliryhmän interventiot ja testit toteutettiin ymmärtääksemme samaan aikaan. Samoin myöhemmässä tutkimuksessa huomattu oireiden paheneminen kaikissa ryhmissä yhdistettiin koronaviruksen aiheuttamaan poikkeustilanteeseen. Koeryhmässä oireiden paheneminen ei kuitenkaan ollut vähäisempää, kuten voisi olettaa. Vaikuttaisi kuitenkin siltä, että ahdistusoireiden pahentuminen ei johtunut tekoälyn käytöstä, vaan kyseessä oli nimenomaan reaktio poikkeukselliseen maailmantilanteeseen. Tutkimusryhmän aikaisemmassa tutkimuksessa tämä ilmeni esimerkiksi niin, että monet osallistujat raportoivat viimeistä terapiakertaa varten tehdyissä päiväkirjamerkinnöissä koronapandemiaan liittyvistä huolista.

Erilaisia tutkimustuloksia voidaan selittää sillä, että mukaan valitut tutkimukset olivat melko heterogeenisiä. Tekoälyä oli hyödynnetty eri tavoin, tutkittavien ikä vaihteli lapsista ikääntyneisiin ja ahdistusta tutkittiin monissa erilaisissa tilanteissa ja eri diagnooseilla. Joissain tutkimuksissa voi jäädä myös epäselväksi, ovatko tulokset nimenomaan seurausta tekoälystä, kun sen osuus interventiossa on pieni tai kun interventioon kuuluu myös osioita, joissa ei käytetä tekoälyä.

Esimerkiksi Petersenin ja kollegoiden tutkimuksen (2017) robottilemmikissä tekoälyä oli vain ”persoonallisuus”. On siis epäselvää, johtuuko tutkimuksessa havaittu ahdistuneisuuden väheneminen nimenomaan tekoälystä vai onko kyseessä pikemminkin suloisen lemmikin vaikutus. Kontrolliryhmän aktiviteetteihin ei kuulunut vastaavaa lemmikkiä. Beidelin ja kollegoiden tutkimuksessa (2021) VR-interventiossa näkyvät avatar-hahmot toimivat tekoälyn avulla, mutta intervention aiemmassa versiossa hahmoja oli hallinnut terapeutti. VR-osiossa käytettiin tekoälyllä toimivia avatar-hahmoja, mutta SET-C interventiossa oli myös osia, joissa ei käytetty tekoälyä lainkaan.

Katsauksessa oli mukana myös tutkimus, jossa tekoälyä ei ollut integroitu itse interventioon (Johnson, 2024). Itsenäinen vuorovaikutus tekoälyn kanssa on voinut olla haasteena erityisesti koehenkilöille, jotka ovat tekoälyn käyttöön tottumattomia. Koehenkilöt saivat myös itse päättää, mitä tekoälytyökalua he käyttivät, joten interventioryhmässä saattoi olla käytössä useita eri tekoälyjä. Koe- ja kontrolliryhmän välillä ei kuitenkaan ollut tilastollisesti merkitseviä eroja.

Otoskoko oli pieni useassa tutkimuksessa (Danieli ym., 2021; Danieli ym., 2022; Sadeh-Sharvit ym., 2023). Näissä tutkimuksissa tutkittavia oli eri ryhmissä 15 tai vähemmän. Pienempi otoskoko voi heikentää tulosten luotettavuutta, koska satunnaisuus vaikuttaa tuloksiin enemmän ja tuloksia on vaikeampi replikoida. Pieni otoskoko ei kuitenkaan näyttäisi näissä tutkimuksissa vaikuttaneen systemaattisesti tutkimustuloksiin – Danielin tutkimusryhmän ensimmäisessä tutkimuksessa

ahdistuneisuus lisääntyi koeryhmässä, kun taas Sadeh-Sharvitin ja kollegoiden tutkimuksissa koe- ja kontrollitilanteiden välillä ei ollut tilastollisesti merkitseviä eroja.

Katsauksemme perusteella näyttäisi siltä, ettei tekoäly auttanut ahdistuneisuuden vähentämisessä merkitsevästi – tarkastelemistamme kymmenestä tutkimuksesta seitsemässä koe- ja kontrollitilanteiden välillä ei ollut tilastollisesti merkitseviä eroja. Kahdessa tutkimuksessa koeryhmässä havaittiin ahdistuneisuuden vähenemistä kontrolliryhmään verrattuna, ja yhdessä ahdistuneisuus paheni tekoälyavusteisen hoidon aikana.

4. Pohdinta

Systemaattisen katsauksemme tavoitteena oli tutkia tekoälyavusteisten hoitomuotojen tehokkuutta ahdistuneisuuden hoidossa tavanomaisiin hoitoihin verrattuna. Lopulliseen tarkasteluun päätyi kymmenen tutkimusta kahden tutkimuksen huonon laadun takia. Tutkimuksissa käytetyt tekoälyteknologiat olivat keskenään erilaisia, samoin ahdistuneisuuden mittarit, mutta näillä seikoilla ei näytä olleen järjestelmällistä vaikutusta tutkimustuloksiin. Katsauksen tulokset vaikuttaisivat olevan yhteneviä aikaisemman tutkimusnäytön kanssa. Pohdimme myös mahdollisia tutkimusaiheeseemme liittyviä jatkotarpeita, ja käsittelemme siihen liittyviä eettisiä kysymyksiä.

4.1 Tuloksiin vaikuttavat tekijät

4.1.1 Heikkolaatuiset tutkimukset

Kahdessa tutkimuksessa (Kamita ym., 2019; Matsumoto ym., 2021) oli huomattavia puutteita, minkä takia emme tarkastelleet niiden tuloksia. Matsumoton ja kollegoiden (2021) tutkimuksen tuloksista ei voi tehdä varteenotettavia päätelmiä muun muassa hyvin pienen lopullisen otoskoon takia. Tutkimuksessa oli huomattavasti pudokkaita, joita ei otettu mukaan lopullisiin analyyseihin. Syynä tähän oli todennäköisesti se, että intervention loppuvaiheen kyselyihin vastaaminen oli vapaaehtoista, koska tutkijat eivät halunneet kuormittaa yövuoroissa työskenteleviä sairaanhoitajia liikaa. Lisäksi tutkittaville annettu ohjeistus vaihteli ryhmittäin, eikä tutkimuksessa ollut toteutettu olennaisia tilastollisia analyyseja, kuten ryhmien vertailua keskenään. Tulokset oli raportoitu epäselvästi, eikä kaikille tekstissä esitetyille tulkinnoille löydy perusteita testituloksista.

Kamitan ja kollegoiden (2019) tutkimuksen tutkimusryhmä on pitkälti sama kuin Matsumoton ja kollegoiden (2021) tutkimuksessa. Kamitan ja kollegoiden tutkimuksen tulosten raportointi on selkeämpää ja esimerkiksi ryhmien väliset p-arvot on esitetty, mutta suhtautuisimme kuitenkin myös tähän tutkimukseen varauksella. Tutkimuksen interventio oli niin ikään hyvin lyhyt; se kesti kokonaisuudessaan vain noin 90 minuuttia, joista 10 minuuttia käytettiin SAT-metodin toteuttamiseen.

Molempien tutkimusten (Kamita ym., 2019; Matsumoto ym., 2021) interventiona käytettiin SAT-menetelmää (Munakata, 2009). Hoitoon liittyvien artikkelien väitteille siitä, mihin hoito perustuu ja millaisia sairauksia sillä voidaan hoitaa (Munakata, 2007), ei löydy riittäviä tieteellisiä perusteita. Munakata esittelee artikkelissaan REI-tekniikan (engl. retrospective evolution imagery), joka on eräs SAT:n menetelmistä. REI:ssä potilas ohjataan kokemaan oman sikiökautensa ja evolutiivisen historiansa vaiheita sekä käsittelemään näiden elämänmuotojen ikäviä muistoja regressiivisen hypnoterapian keinoin, mikä muistuttaa jälleensyntymän ajatukseen perustuvaa regressioterapiaa eli menneet elämät -terapiaa. Asiantuntijoiden mukaan regressioterapia on hoitomuotona epäeettinen, koska sen toimivuudesta ei ole tutkimusnäyttöä ja hoidon aikana syntyvät valemuistot voivat olla haitaksi potilaalle (Andrade, 2017). Kamita ja Matsumoto tutkimusryhmineen eivät kuitenkaan ilmeisesti harjoittaneet REI-tekniikkaa SAT-interventioiden yhteydessä, mutta tämä tekniikka osana SAT-metodia horjuttaa koko intervention luotettavuutta.

Munakata esittää lisäksi artikkelissaan (2007), että SAT-menetelmän avulla on mahdollista hoitaa syöpää, sillä menetelmällä voidaan lisätä kehon lymfosyyttien ja antionkogeenien *p53* ja *RB* määrää. Huomionarvoista on myös, että Munakata on yksi kirjoittajista sekä Kamitan että Matsumoton tutkimusryhmien artikkeleissa. Pohdimme myös sitä, voiko osa tutkimuksissa havaitsemistamme heikkouksista johtua käänkövirheistä, mutta tutkimukset vaikuttavat kieliseikoista huolimatta niin epäluotettavilta, ettemme anna niiden tuloksille painoarvoa tämän katsauksen kokonaistuloksissa.

4.1.2 Muut tuloksiin vaikuttavat tekijät

Osassa katsauksen tutkimuksista oli paljon pudokkaita. Esimerkiksi Sadeh-Sharvitin ja kollegoiden (2023) tutkimuksessa eri syistä keskeyttäneitä tai osaan kyselyistä vastaamatta jättäneitä oli 53 % (25/47) alkuperäisistä osallistujista. Pudokkaat voivat aiheuttaa vinoumaa tuloksiin, tosin se ei ole vääjäämätöntä (Bell ym., 2013). Esimerkiksi jos eri interventoryhmissä on eri määrä pudokkaita tai erilaisia syitä tutkimuksesta poisjättäytymiseen, on tähän aiheellista kiinnittää huomiota (Moher ym., 2010). Sadeh-Sharvitin ja kollegoiden tutkimuksessa pudokkaita oli kuitenkin suurin piirtein yhtä paljon eri interventoryhmissä. Pudokkaiden määrä ei siis tässä tapauksessa näyttäisi vaikuttavan tuloksiin muuten kuin pienen otoskoon kautta.

Myös oireettomat populaatiot voivat rajoittaa tilastollisesti merkitsevien tulosten saamista. Suhawardyn ja kumppaneiden (2023) tutkittavat eivät alkumittauksessa ylittäneet kliinisen ahdistuneisuuden rajaa, eikä tutkimuksella ollut heidän mukaansa riittävää voimaa erojen tunnistamiseen subkliinisessä populaatiossa.

Lähes kaikissa tutkimuksissa otoksen sukupuolijakauma oli painottunut niin, että suurin osa otoksesta oli naisia. Poikkeuksena tähän oli Russon ja kollegoiden (2024) tutkimus, jossa otos koostui tasan puoliksi miehistä ja puoliksi naisista. Suhawardyn ja kollegoiden (2023) tutkimuksen kaikki

koehenkilöt olivat vastikään synnyttäneitä naisia. Tiedetään, että ahdistuneisuushäiriöt ovat yleisempiä naisilla kuin miehillä (esim. Munk-Jorgensen ym., 2006). Toisaalta saattaa myös olla, että naisia on helpompi rekrytoida tutkimuksiin koehenkilöiksi. Voi kuitenkin pohtia, voidaanko hoitomenetelmien olettaa sopivan kaikille sukupuolille, jos tutkittavista suurin osa on naisia.

On myös mahdollista, että tietyt ahdistuneisuudelle tyypilliset piirteet vaikuttavat siihen, millä tavalla tutkittavat käyttävät tekoälyä ja miten tekoälyn käyttö vaikuttaa heidän mielenterveyteensä. Johnson ja kollegat (2024) pohtivat tutkimuksessaan, että perfektionismi saattaa vaikuttaa tekoälyn käyttöön esimerkiksi siten, että perfektionisti voi vältellä tekoälyn käyttöä, jos hän kokee, että se voidaan tulkita merkiksi hänen omasta osaamattomuudestaan ja riittämättömydestään. Tekoälyn käyttö voi heidän mukaansa myös aiheuttaa suorituspainetta käyttäjille, jotka ovat jo valmiiksi taipuvaisia niiden kokemiseen.

Keskusteluagenttien yksi nykyongelma on, ettei muilla kuin englannin kielellä ole välttämättä riittävästi dataa tekoälyn kouluttamiseen, mikä voi vaikuttaa tekoälyn tehokkuuteen (Guo ym., 2024; Zygadlo, 2021). Danielin ja kumppaneiden (2021, 2022) artikkeleissa ei suoraan kerrota, millä kielellä interventiot toteutettiin, mutta tutkittavat olivat italialaisia, joten oletamme, että keskusteluagentti on myös toiminut italian kielellä. Vaikka tämän tutkimuksen tulosten perusteella keskusteluagentin kieli ei näyttäisi erityisesti vaikuttavan tuloksiin, asiaan on hyvä kiinnittää huomiota keskusteluagentti-interventioita tutkittaessa.

Näiden seikkojen lisäksi pohdimme tutkijan sidonnaisuuksien mahdollista vaikutusta tuloksiin. Gleasonin ja kollegoiden (2025) tutkimuksen oli rahoittanut Woebot Health, jonka bottia käytettiin tutkimuksessa. Sadeh-Sharvitin tutkimusryhmästä (2023) kaksi kirjoittajaa puolestaan työskenteli Eleos Health-yritykselle. Näiden tutkimusten tulokset eivät eronneet systemaattisesti muista katsaukseen valituista tutkimuksista, mutta sidonnaisuuksien mahdollista vaikutusta tuloksiin ei tule sivuuttaa.

4.2 Tutkimuksen vahvuudet ja rajoitukset

Tämä katsaus on ensimmäinen systemaattinen katsaus, jossa tekoälyavusteisten hoitojen tehokkuutta on pyritty tutkimaan aktiivisilla kontroleilla. Systemaattinen tiedonhaku kahdesta tietokannasta tuotti kattavasti tuloksia. Käyttämämme tietokantakohtaiset hakulausekkeet olivat niin ikään perusteelliset, ja tarkkojen sisäänotto- ja poissulkukriteerien avulla löysimme sopivia artikkeleita. Katsauksen luotettavuutta paransi kaksi kirjoittajaa sekä toteuttamamme sokkoutus aineiston hankintavaiheessa.

Katsaukseen haluttiin ottaa mukaan tutkimuksia, joissa kontrollitilanne on aktiivinen. Lopullisessa aineistossa kaikki kontrollit olivat aktiivisia lukuun ottamatta Johnsonin ja kollegoiden (2024) tutkimusta, jossa myös koetilanne oli tekoälyavusteinen omahoito ja kontrollitilanne pelkkä omahoito. Näin pystyimme tutkimaan nimenomaan tekoälyn vaikutusta hoidon tehokkuuteen.

Tutkimuksen aineisto oli varsin heterogeeninen, mikä voi olla sekä tutkimuksen vahvuus että sen rajoitus. Heterogeenisillä aineistoilla saadut tulokset voivat päteä monissa erilaisissa tilanteissa, mutta toisaalta on tarpeen arvioida, voiko hyvin erilaisista populaatioista tehdä johtopäätöksiä. Tämän katsauksen tutkimuksista suurimmassa osassa ryhmien välillä ei ollut merkitseviä eroja, joten heterogeenisyyden takia jää vielä epäselväksi, voiko näitä tuloksia yleistää. On mahdollista, että jokin tekoälymenetelmistä voisi toimia tietyissä tilanteissa.

Tutkimuksen vahvuutena oli myös tutkimusten ajantasaisuus – yhdeksän artikkelia kymmenestä oli julkaistu kuluneen neljän vuoden aikana. Tutkittavien ikähaarukka oli niin ikään laaja. Tutkimuksista kuusi sijoittui Yhdysvaltoihin, kolme Italiaan ja yksi Australiaan, mikä puolestaan saattaa heikentää tulosten yleistettävyyttä länsimaiden ulkopuolella.

Tekoäly käsitteenä on hyvin laaja, minkä takia tutkimuksessa oli haastavaa määritellä, mitkä interventiot käyttävät tekoälyä. Opimme tutkimusprosessin aikana esimerkiksi, että chattibotit voivat toimia myös ilman tekoälyä, minkä takia jouduimme kokotekstivaiheessa karsimaan vielä tutkimuksia. Katsaukseen valikoitui lopulta varsin moninainen joukko erilaisia tekoälyinterventioita, joten voi olla tarpeellista miettiä, voidaanko tuloksista tehdä suoraan johtopäätöksiä.

Tekoäly on hyvin ajankohtainen aihe, ja sitä aletaan soveltaa myös terveydenhuollossa yhä enemmän. On siis tarpeen koota tietoa tekoälyä hyödyntävien hoitojen tehokkuudesta.

4.3 Tulosten suhteutuminen aikaisempaan tutkimukseen

Tämän katsauksen perusteella tekoälyn hyödyntäminen ei näyttäisi vaikuttavan ahdistuneisuuden hoitomenetelmän tehokkuuteen. Aikaisempaa tutkimusnäyttöä, jossa tekoälyinterventiota verrataan aktiiviseen kontrollitilanteeseen, on niukasti. Vaikuttaa kuitenkin siltä, että katsauksen tulokset ovat yhteneviä aikaisemman tutkimusnäytön kanssa. Suurimmassa osassa katsauksen tutkimuksista koe- ja kontrolliryhmien tulokset eivät eronneet merkitsevästi toisistaan, mikä tukee Dasin ja Gavaden (2024) havaintoa siitä, että tekoälyavusteisten hoitojen hyödyt liittyvät nimenomaan hoidon saatavuuteen, yksilöllisyyteen ja hallinnollisten prosessien tukemiseen eivätkä niinkään ahdistuneisuuden vähentämiseen.

4.4 Tutkimuksen merkitys

Vaikuttaisi siltä, että tekoäly ei suoranaisesti lisää ahdistuneisuuden hoidon tehokkuutta, mutta sen käytöllä voidaan saada muita hyötyjä. Tekoäly voi olla varteenotettava hoitomuoto erityisesti saavutettavuuden ja terapeuttien resurssien säästämisen näkökulmasta. Vaikka tekoälyavusteiset hoitomuodot vaikuttavat olevan yhtä tehokkaita verrattuna samankaltaisiin interventioihin, joissa ei käytetä tekoälyä, voi tekoälyinterventioilla olla kuitenkin kliinistä merkittävyyttä. Tekoälyinterventioiden käytöstä herää joka tapauksessa erilaisia eettisiä ja tietosuojaan liittyviä kysymyksiä.

4.4.1 Eettiset kysymykset

Tekoälyn hyödyntäminen etenkin terapiakäytössä herättää kysymyksiä hoidon eettisyydestä. Vaikka tekoäly olisi ammattilaisten kouluttama, ei sillä kuitenkaan ole aitoa ammattitaitoa, empatiaa tai kykyä ymmärtää ihmistunteiden nyansseja.

Yksi suurten kielimallien ongelmista on ns. *hallusinointi* (engl. hallucination), jolla tarkoitetaan mallin tuottamaa väärää tietoa. Suurten kielimallien hallusinoinnin on osoitettu olevan väistämätöntä (Xu ym., 2024). Ongelma piilee siinä, että tekoälyllä ei ole metakognitiota – ihmisterapeutti voi todeta, ettei tiedä jotakin, mutta tekoäly voi tuottaa väärää tietoa kyseenalaistamatta sen todenperäisyyttä.

Tekoälyn terapiakäytössä on ensiarvoista huolehtia myös tietosuojasta (Alfano ym., 2024). Potilastietojen tulee olla suojattu asianmukaisesti, jotta tietomurtojen riskiä ja niistä aiheutuvaa haittaa potilaille voidaan vähentää. Tekoölyyn liittyvä lainsäädäntö on vielä osin alkutekijöissään, minkä takia yksi huolenaihe on, mihin tekoälyterapiassa käytetty data tallennetaan ja kuka siitä on vastuussa. Olemassa on myös riski datan jakamisesta ja myymisestä eteenpäin.

Tekoäly näyttäisi kykenevän tunnistamaan tiettyjä kriisitilanteita, kuten itsemurhariskin (Khan & Javed, 2022; Menon & Vijayakumar, 2023). Systemaattisen katsauksemme tutkimuksista Gleasonin ja kumppaneiden (2025) tekoälyinterventiossa käytettiin niin sanottua ”turvaverkkoa”, jossa tekoäly tunnisti keskustelusta mahdollisia kriisitilanteeseen viittaavia sanoja, ja hälyytti asiasta potilasta hoitavalle ammattilaiselle. Tutkimuksessa ”turvaverkko” onnistui tekemään hälytykset kriisitilanteista ja teki tämän lisäksi yhden väärän hälytyksen neutraalin sanan virheellisen tulkinnan takia. Vaikka tekoäly voikin tunnistaa kriisitilanteita tehokkaammin kuin ihminen, säilyy vastuu hoidosta terapeuteilla eettisistä ja laillisista syistä (Alfano ym., 2024). Tekoäly voi olla nopea arvioissaan, mutta sillä ei ole samanlaista ammattiosaamista, moraalialia tai ymmärrystä yksittäistapauksista kuin ihmisterapeutilla, mikä voi johtaa potilaan tilanteen vääränlaiseen tulkintaan.

Vaikka tekoälyä voisi pitää objektiivisena, on se myös taipuvainen erilaisille vinoumille (Timmons ym., 2022), sillä tekoälytyökalut ovat ihmisten tuotosta ja niiden kouluttamiseen on käytetty ihmisten luomaa ja valikoimaa sisältöä. Tekoäly voi siis siten toistaa ja jopa vahvistaa koulutusdatan mahdollisesti sisältämiä vinoumia. Tällaisista tekoälyn vinoumista erilaisissa terveydenhuollon sovelluksissa on tutkimusnäyttöä (ks. esim. Kamulegeya ym., 2023; Obermeyer ym., 2019). Mielenterveyden hoidossa olisi tärkeää huomioida, että tunteiden käsitteellistäminen ja ilmaisu vaihtelevat huomattavasti eri sukupuolilla ja kulttuureilla (Office of the Surgeon General, 2001). Esimerkiksi luonnollisen kielen käsittelyjärjestelmien suorituskyvyn on osoitettu vaihtelevan ihmisen

sukupuolen, uskonnon, seksuaalisuuden, etnisyyden ja iän perusteella (Straw & Callison-Burch, 2020).

4.4.2 Jatkotutkimus- ja sovellusehdotukset

Terapeutin ja asiakkaan välisellä yhteistyösuhteella eli terapeuttisella allianssilla on yhteys psykoterapian tehokkuuteen (ks. esim. Horvath ym., 2011), minkä takia sitä olisi mielenkiintoista tutkia myös tekoälyterapian näkökulmasta. Voiko tekoälyn kanssa muodostaa samanlaista suhdetta kuin ihmisen kanssa, ja miten ihmiskontaktin puuttuminen voi vaikuttaa terapian tehokkuuteen, kun luottamuksellisen ja toimivan terapiasuhteen rakentuminen on yksi hoidon kulmakivistä?

Suharwardyn ja kumppaneiden (2023) tutkimuksessa saatiin näyttöä terapeutin allianssin ja kiintymissuhteen syntymisestä ihmisen ja Woebotin välille. Chattibottia käyttäneet osoittivat positiivisen terapeutin allianssin syntymistä botin kanssa WAI-SR:llä (Working Alliance Inventory-Short Revised) mitattuna. Vastaavia tuloksia on saatu myös esimerkiksi päihteiden väärinkäytön vähentämiseen suunnitellulla chattibotti-interventiolla (Prochaska ym., 2021). Kiintymyssuhteita, joita tutkittavat ovat kehittäneet chattibottiin, on kuvailtu myös laadullisesti (esim. Fitzpatrick ym., 2017; Fulmer ym., 2018). Ilmiön voi katsoa peilaavan perinteisissä terapiasuhteissa ilmenevää transferenssia (Joseph & Babu, 2024).

Tutkimusta transferenssista tekoälyterapiassa on vielä vähäisesti (ks. Holohan & Fiske, 2021; Joseph & Babu, 2024). Transferenssilla tarkoitetaan ilmiötä, jossa tärkeissä ihmissuhteissa koettuja tunteita kohdistetaan tiedostamatta muihin ihmissuhteisiin, kuten terapiasuhteeseen (Freud, 1912; Levy & Scala, 2012) – vastatransferenssi puolestaan tarkoittaa terapeutin tunteiden kohdistamista asiakkaaseen. Transferenssi on kiinteä osa terapiasuhdetta, ja se voi vahvistaa terapeutista allianssia ja auttaa vaikeiden aiheiden käsittelemisessä, mutta myös toimia hoidon esteenä, jos asiakas esimerkiksi kokee, ettei voi keskustella tietyistä aiheista terapeutin kanssa. Vastatransferenssi voi olla vahingollista terapiasuhteelle (Laaksonen-Balk, 2023), mutta toisaalta voi auttaa terapeutia ymmärtämään asiakkaan sisäistä maailmaa ja niin ikään vahvistaa terapiasuhdetta (Quamme, 2022). Transferenssi voi siis olla sekä hyödyksi että haitaksi terapiasuhteelle, minkä takia sen ilmenemistä tekoälyterapiassa olisi kiintoisaa tutkia etenkin tekoälytyökalujen kehittyessä ihmismäisemmiksi.

Myös erilaisten psyykkisten häiriöiden erityispiirteiden ymmärtäminen on transferenssi-ilmiön ohella tärkeää, kun tekoälyinterventioita kehitetään. Keskustelu tekoälyterapeutin kanssa voi auttaa ihmisiä, jotka epäröivät avun hakemista tuomitsemisen ja leimautumisen pelon tai ennakkoluulojen vuoksi (Khawaja & Bélisle-Pipon, 2023), ja siten toimia esimerkiksi sosiaalisesta ahdistuneisuudesta kärsiville. Toisaalta tekoälyterapeutin helposti saavutettava tuki voi aiheuttaa riippuvuutta, mikä voi Josephin ja Babun (2024) mukaan kapeuttaa reaali maailman vuorovaikutussuhteita ja nostaa kynnystä osallistua ihmisterapeutin toteuttamaan hoitoon. Lisäksi vuorovaikutus tekoälyn kanssa on

ennustettavaa ja johdonmukaista, mikä voi toisaalta vähentää ahdistuksen laukaisevia tekijöitä ja luoda uhattoman terapiaympäristön, muttei kuitenkaan ole esimerkiksi altistusterapiassa tarkoituksenmukaista. Tutkijoiden mukaan tekoälyterapiassa on myös muihin psyykkisiin häiriöihin, kuten persoonallisuushäiriöihin, pakko-oireiseen häiriöön, skitsofreniaan ja kaksisuuntaiseen mielialahäiriöön liittyviä niille ominaisia transferenssin sudenkuoppia.

Katsauksemme tutkimuksissa hyödynnetyt tekoälymallit olivat pitkälti tekstiä tuottavia ja sitä analysoivia malleja. Siksi olisi mielenkiintoista tutkia myös tekoälypohjaisten äänimallien hyödyntämistä terapiatarkoituksiin. Tässä yhteydessä voitaisiin käyttää esimerkiksi ns. *puhe puheeksi* (engl. speech-to-speech) -työkaluja, jolloin käyttäjä voi puhua tekoälyterapeutille, joka puolestaan vastaa puhuen. Yhdysvaltalainen startup-yritys Sesame (2025) on vastikään julkaissut avoimen lähdekoodin *keskustelevan puhemallin* (engl. Conversational Speech Model). Sesamen puhemallilla voidaan tuottaa hyvin luonnollisen kuuloista keskustelua, ja malli on hyvä tunnistamaan ihmisen puheen sävyjä ja tahtia (Buntz, 2025). Arvelemme, että tällaisella äänimallilla toteutettu tekoälyterapia voisi muistuttaa enemmän perinteistä, kasvoitusten tapahtuvaa terapiaa verrattuna tekstipohjaiseen chattibottiin, ja siten olla helpommin lähestyttävää ja mahdollisesti hyväksyttävämpää sekä luoda paremman pohjan terapeuttisen yhteistyösuhteen syntymiselle.

4.5 Yhteenveto

Tekoälyn käyttö ahdistuneisuuden hoidossa on ajankohtainen ilmiö. Tämän katsauksen perusteella tekoälyn lisäämisellä hoitoon ei näyttäisi olevan vaikutusta ahdistuneisuuden vähentämisessä. Tekoälyn hyödyntämisellä voi olla kuitenkin kliinistä merkittävyyttä, jos sen käyttö mahdollistaa muita hyötyjä ahdistuksen hoidossa. Tekoälyä käyttämällä voitaneen säästää rahaa ja sekä asiakkaiden että terapeuttien aikaa. Esimerkiksi chattibotilla toteutettu interventio on helposti saavutettavissa alueilla, joissa on pulaa terapeuteista. Tekoäly voi tehostaa terapeuttien työtä esimerkiksi hoitamalla hallinnollisia tehtäviä ja tekemällä yhteenvetoja terapiakäynneistä. Näistä hyödyistä huolimatta on kuitenkin ensiarvoista seurata tarkasti, että tekoälyavusteiset hoitomuodot toteutetaan eettisesti ja potilaan oikeudet huomioiden. Myös eri psyykkisten häiriöiden erityispiirteet tulisi ottaa huomioon interventioiden suunnittelussa.

Tekoälyllä ei voida korvata ihmisläheistä otetta terapiaan, eikä sen ammattitaito vastaa ihmisen ammattitaitoa. Tekoäly voi monessa tapauksessa olla asianmukainen lisä jonkin terapian osa-alueen tehostamisessa, mutta ihmisterapeuttien tarjoamaa terapiaa ei sillä pystytä täysin korvaamaan.

Lähteet

- Alfano, L., Malcotti, I., & Ciliberti, R. (2024). Psychotherapy, artificial intelligence and adolescents: ethical aspects. *Journal of Preventive Medicine and Hygiene*, 64(4), E438–E442. <https://doi.org/10.15167/2421-4248/jpmh2023.64.4.3135>
- American Psychiatric Association. (1994). *DSM-IV: Diagnostic and Statistical Manual of Mental Disorders* (4. painos). <https://doi.org/10.1001/jama.1994.03520100096046>
- Andrade G. (2017). Is past life regression therapy ethical? *Journal of Medical Ethics and History of Medicine*, 10, 11.
- Antony, M. M., Purdon, C. L., Huta, V., & Swinson, R. P. (1998). Dimensions of perfectionism across the anxiety disorders. *Behaviour Research and Therapy*, 36(12), 1143–1154. [https://doi.org/10.1016/S0005-7967\(98\)00083-7](https://doi.org/10.1016/S0005-7967(98)00083-7)
- Beck, J. S. (2020). *Cognitive Behavior Therapy: Basics and beyond*. (3. painos). Guilford Publications.
- Beg, M. J., Verma, M., M., V. C., & Verma, M. K. (2024). Artificial Intelligence for Psychotherapy: A review of the current state and Future Directions. *Indian Journal of Psychological Medicine*. <https://doi.org/10.1177/02537176241260819>
- Beidel, D. C., Turner, S. M., & Morris, T. L. (1995). A new inventory to assess childhood social anxiety and phobia: The Social Phobia and Anxiety Inventory for Children. *Psychological Assessment*, 7(1), 73–79. <https://doi.org/10.1037/1040-3590.7.1.73>
- Beidel, D.C., Turner, S.M., & Morris, T.L. (2004). *Social Effectiveness Therapy for Children and Adolescents (SET-C)*. Multi-Health Systems, Inc.
- *Beidel, D. C., Tuerk, P. W., Spitalnick, J., Bowers, C. A., & Morrison, K. (2021). Treating Childhood Social Anxiety Disorder with Virtual Environments and Serious Games: A Randomized Trial. *Behavior Therapy*, 52(6), 1351–1363. <https://doi.org/10.1016/j.beth.2021.03.003>
- Bell, M. L., Kenward, M. G., Fairclough, D. L., & Horton, N. J. (2013). Differential dropout and bias in randomised controlled trials: when it matters and when it may not. *BMJ*, 346, e8668. <https://doi.org/10.1136/bmj.e8668>
- Bergström, J., Andersson, G., Ljótsson, B., Rück, C., Andréewitch, S., Karlsson, A., Carlbring, P., Andersson, E., & Lindefors, N. (2010). Internet-versus group-administered cognitive behaviour therapy for panic disorder in a psychiatric setting: a randomised trial. *BMC Psychiatry*, 10, 54. <https://doi.org/10.1186/1471-244X-10-54>

Buntz, B. (14.3.2025). *The R&D story behind Sesame AI, the startup that just open-sourced its groundbreaking voice generation model*. Haettu 20.4.2025 osoitteesta <https://www.rdworldonline.com/the-rd-story-behind-sesame-ai-the-startup-that-just-open-sourced-its-groundbreaking-voice-generation-model/>

*Danieli, M., Ciulli, T., Mousavi, S. M., & Riccardi, G. (2021). A Conversational Artificial Intelligence Agent for a Mental Health Care App: Evaluation Study of Its Participatory Design. *JMIR Formative Research*, 5(12), e30053. <https://doi.org/10.2196/30053>

*Danieli, M., Ciulli, T., Mousavi, S. M., Silvestri, G., Barbato, S., Di Natale, L., & Riccardi, G. (2022). Assessing the Impact of Conversational Artificial Intelligence in the Treatment of Stress and Anxiety in Aging Adults: Randomized Controlled Trial. *JMIR Mental Health*, 9(9), e38067. <https://doi.org/10.2196/38067>

Das, K. P., & Gavade, P. (2024). A review on the efficacy of Artificial Intelligence for managing anxiety disorders. *Frontiers in Artificial Intelligence*, 7. <https://doi.org/10.3389/frai.2024.1435895>

Ellis, A. (1962). *Reason and emotion in psychotherapy*. Lyle Stuart.

Farzan, M., Ebrahimi, H., Pourali, M., & Sabeti, F. (2025). Artificial Intelligence-Powered Cognitive Behavioral Therapy Chatbots, a Systematic Review. *Iranian Journal of Psychiatry*, 20(1), 102–110. <https://doi.org/10.18502/ijps.v20i1.17395>

Fenn, K., & Byrne, M. (2013). The key principles of cognitive behavioural therapy. *InnovAiT*, 6(9), 579–585. <https://doi.org/10.1177/1755738012471029>

Fitzpatrick, K.K., Darcy, A., & Vierhile, M. (2017). Delivering Cognitive Behavior Therapy to Young Adults with Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Mental Health*, 4(2), e19. <https://doi.org/10.2196/mental.7785>

Freud, S. (1912). The dynamics of transference. Teoksessa Strachey, J. (toim.). *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XII (1911–1913): The Case of Schreber, Papers on Technique and Other Works*, 97–108.

Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled Trial. *JMIR Mental Health*, 5(4), e64. <https://doi.org/10.2196/mental.9782>

Gale, C., & Davidson, O. (2007). Generalised anxiety disorder. *BMJ*, 334(7593), 579–581. <https://doi.org/10.1136/bmj.39133.559282.BE>

- *Gleason, M. M., Flom, M., Rapoport, S., Williams, A., Birch, A., Wells, N. K., Forman-Hoffman, V., & Robinson, A. (2025). A Relational Agent Intervention for Adolescents Seeking Mental Health Treatment: Outcomes from a Randomized Controlled Trial Within a Children's Outpatient Hospital. *JAACAP Open*. <https://doi.org/10.1016/j.jaacop.2025.02.002>
- Gould, R. A., Ott, M. W., & Pollack, M. H. (1995). A meta-analysis of treatment outcome for panic disorder. *Clinical Psychology Review*, *15*(8), 819-844. [https://doi.org/10.1016/0272-7358\(95\)00048-8](https://doi.org/10.1016/0272-7358(95)00048-8)
- Guo, Z., Lai, A., Thygesen, J. H., Farrington, J., Keen, T., & Li, K. (2024). Large language models for Mental Health Applications: Systematic Review. *JMIR Mental Health*, *11*. <https://doi.org/10.2196/57400>
- Guy, W. (1976). CGI Clinical Global Impressions. *ECDEU assessment manual for psychopharmacology*. (Uudistettu painos, 217–222). U.S. Department of Health and Human Services, Public Health Service, Alcohol Drug Abuse and Mental Health Administration, National Institute of Mental Health Psychopharmacology Research Branch.
- Haravuori, H. (2020). Lyhytkestoisen psykoterapiatekniikan muoto Eye Movement Desensitization and Reprocessing (EMDR). *Duodecim Käypä hoito*. www.kaypahoito.fi
- Helm, J. M., Swiergosz, A. M., Haeberle, H. S., Karnuta, J. M., Schaffer, J. L., Krebs, V. E., Spitzer, A.I., & Ramkumar, P. N. (2020). Machine learning and artificial intelligence: definitions, applications, and future directions. *Current Reviews in Musculoskeletal Medicine*, *13*, 69–76. <https://doi.org/10.1007/s12178-020-09600-8>
- Higa, C. K., Fernandez, S. N., Nakamura, B. J., Chorpita, B. F., & Daleiden, E. L. (2006). Parental assessment of childhood social phobia: Psychometric properties of the Social Phobia and Anxiety Inventory for Children–Parent Report. *Journal of Clinical Child and Adolescent Psychology*, *35*(4), 590–597. https://doi.org/10.1207/s15374424jccp3504_11
- Holohan, M., & Fiske, A. (2021). “Like I’m Talking to a Real Person”: Exploring the Meaning of Transference for the Use and Design of AI-Based Applications in Psychotherapy. *Frontiers in Psychology*, *12*, 720476. <https://doi.org/10.3389/fpsyg.2021.720476>
- Horowitz, M. Wilner, N., & Alvarez, W. (1979). Impact of Event Scale: A measure of subjective stress. *Psychosomatic Medicine*, *41*, 209–218. <https://doi.org/10.1097/00006842-197905000-00004>
- Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy*, *48*(1), 9–16. <https://doi.org/10.1037/a0022186>

- Jiang, M., Zhao, Q., Li, J., Wang, F., He, T., Cheng, X., Yang, B. X., Ho, G. W. K., & Fu, G. (2024). A Generic Review of Integrating Artificial Intelligence in Cognitive Behavioral Therapy. *arXiv*. <https://doi.org/10.48550/arXiv.2407.19422>
- *Johnson, C., Egan, S. J., Carlbring, P., Shafran, R., & Wade, T. D. (2024). Artificial intelligence as a virtual coach in a cognitive behavioural intervention for perfectionism in young people: A randomised feasibility trial. *Internet Interventions*, 38, 100795. <https://doi.org/10.1016/j.invent.2024.100795>
- Joseph, A. P., & Babu, A. (2024). Transference and the psychological interplay in AI-enhanced mental healthcare. *Frontiers in Psychiatry*, 15, 1460469. <https://doi.org/10.3389/fpsyt.2024.1460469>
- *Kamita, T., Ito, T., Matsumoto, A., Munakata, T., & Inoue, T. (2019). A Chatbot System for Mental Healthcare Based on SAT Counseling Method. *Mobile Information Systems*, 2019, 1–11. <https://doi.org/10.1155/2019/9517321>
- Kamulegeya, L., Bwanika, J., Okello, M., Rusoke, D., Nassiwa, F., Lubega, W., Musunguzi, D., & Börve, A. (2023). Using artificial intelligence on dermatology conditions in Uganda: A case for diversity in training data sets for machine learning. *African Health Sciences*, 23(2), 753–63. <https://doi.org/10.4314/ahs.v23i2.86>
- Khan, N. Z., & Javed, M. (2022). Use of Artificial Intelligence-Based Strategies for Assessing Suicidal Behavior and Mental Illness: A Literature Review. *Cureus* 14(7), e27225. <https://doi.org/10.7759/cureus.27225>
- Khawaja, Z., & Bélisle-Pipon, J-C. (2023). Your robot therapist is not your therapist: understanding the role of AI-powered mental health chatbots. *Frontiers in Digital Health*, 5, 1278186. <https://doi.org/10.3389/fdgth.2023.1278186>
- Koponen, H., Laukkala, T., Lappalainen, J., Lehti, V., Leinonen, E., Lindeman, S., Marttunen, M., Niemelä, S., Purhonen, M., Räsänen, K., Valkonen-Korhonen, M., & Viikki, M. (2024). Ahdistuneisuushäiriöt. Käypä hoito -suositus. *Duodecim Käypä hoito*. <https://www.kaypahoito.fi/>
- Kwak, Y. T., Yang, Y., & Koo, M. S. (2017). Anxiety in dementia. *Dementia and Neurocognitive Disorders*, 16(2), 33. <https://doi.org/10.12779/dnd.2017.16.2.33>
- Laaksonen-Balk, T. (25.4.2023). *Miten kohtaan mielenterveyshäiriöistä oireilevan?* [PowerPoint-diat]. Pohjois-Savon hyvinvointialue.

<https://pshyvinvoitorialue.fi/documents/594193/723710/Miten+kohtaan+mielenterveysh%C3%A4iri%C3%B6ist%C3%A4+oireilevan+Teijamari+Laasonen-Balk+25.4.2023.pdf/b5cd0471-317c-6c47-7456-96a1c37f2334?t=1682658232541>

*Leo, A. J., Schuelke, M. J., Hunt, D. M., Miller, J. P., Areán, P. A., & Cheng, A. L. (2022). Digital Mental Health Intervention Plus Usual Care Compared with Usual Care Only and Usual Care Plus In-Person Psychological Counseling for Orthopedic Patients with Symptoms of Depression or Anxiety: Cohort Study. *JMIR Formative Research*, 6(5), e36203.

<https://doi.org/10.2196/36203>

Levy, K. N., & Scala, J. (2012). Transference, transference interpretations, and transference-focused psychotherapies. *Psychotherapy*, 49(3), 391. <https://doi.org/10.1037/a0029371>

Lim, S. M., Shiau, C. W., Cheng, L. J., & Lau, Y. (2022). Chatbot-delivered psychotherapy for adults with depressive and anxiety symptoms: A systematic review and meta-regression. *Behavior Therapy*, 53(2), 334–347. <https://doi.org/10.1016/j.beth.2021.09.007>

*Matsumoto, A., Kamita, T., Tawaratsumida, Y., Nakamura, A., Fukuchimoto, H., Mitamura, Y., Suzuki, H., Munakata, T., & Inoue, T. (2021). Combined Use of Virtual Reality and a Chatbot Reduces Emotional Stress More Than Using Them Separately. *Journal of Universal Computer Science*, 27(12). <https://doi.org/10.3897/jucs.77237>

Menon, V., & Vijayakumar, L. (2023). Artificial intelligence-based approaches for suicide prediction: Hope or hype? *Asian Journal of Psychiatry*, 88, 103728.

<https://doi.org/10.1016/j.ajp.2023.103728>

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., & Devereaux, P. J. (2010). CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*, 340, c869. <https://doi.org/10.1136/bmj.c869>

Munakata, T. (2007). Building SAT Imagery Therapy to Activate Anti-Cancer Genes and Immunologic Function for Cancer Treatment. *International Journal of Structured Association Technique*, 1, 3–36.

Munakata, T. (2009). Reconstructing life and society with SAT therapy: foundations of the new generation CBT. *International Journal of Structured Association Technique*, 3, 35–60.

Munk-Jorgensen, P., Allgulander, C., Dahl, A. A., Foldager, L., Holm, M., Rasmussen, I., Virta, A., Huuhtanen, M.-T., & Wittchen, H.-U. (2006). Prevalence of generalized anxiety disorder in general practice in Denmark, Finland, Norway, and Sweden. *Psychiatric Services*, 57(12), 1738–1744. <https://doi.org/10.1176/appi.ps.57.12.1738>

- Nurmilaakso, T., & Åström, M. (28.04.2024). Nuorten mielenterveysongelmat ovat räjähtämässä käsiin, mutta ilmiössä on myös toivo. *Yle Akuutti*. <https://yle.fi/aihe/a/20-10006638>
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453. <https://doi.org/10.1126/science.aax2342>
- OECD & European Observatory on Health Systems and Policies. (2023). Finland: Country Health Profile 2023. State of Health in the EU. *OECD Publishing*. <https://doi.org/10.1787/e7af1b4d-en>
- Office of the Surgeon General. (2001). Culture counts: The influence of culture and society on mental health. Teoksessa *Mental Health: Culture, Race, and Ethnicity: A Supplement to Mental Health: A Report of the Surgeon General*. Substance Abuse and Mental Health Services Administration (US). <https://www.ncbi.nlm.nih.gov/books/NBK44249/>
- Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan — a web and mobile app for systematic reviews. *Systematic Reviews*, 5(1), 1–10. <https://doi.org/10.1186/s13643-016-0384-4>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The Prisma 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- *Petersen, S., Houston, S., Qin, H., Tague, C., & Studley, J. (2017). The Utilization of Robotic Pets in Dementia Care. *Journal of Alzheimer's Disease*, 55(2), 569–574. <https://doi.org/10.3233/JAD-160703>
- Prochaska, J. J., Vogel, E. A., Chieng, A., Baiocchi, M., Maglalang, D. D., Pajarito, S., Weingardt, K. R., Darcy, A., & Robinson, A. (2021). A randomized controlled trial of a therapeutic relational agent for reducing substance misuse during the COVID-19 pandemic. *Drug and Alcohol Dependence*, 227, 108986. <https://doi.org/10.1016/j.drugalcdep.2021.108986>
- Quamme, J. (2.5.2022). Countertransference and the Therapeutic Relationship. Haettu 20.4.2025 osoitteesta <https://counselormagazine.com/article/countertransference-and-the-therapeutic-relationship/>
- Rovasalo, A., & Eerola, H. (31.1.2023). Ahdistus ja ahdistuneisuus. Lääkärikirja Duodecim. *Kustannus Oy Duodecim*. Haettu 10.4.2025 osoitteesta www.terveyskirjasto.fi

- *Russo, S., Fiani, F., & Napoli, C. (2024). Remote Eye Movement Desensitization and Reprocessing Treatment of Long-COVID- and Post-COVID-Related Traumatic Disorders: An Innovative Approach. *Brain Sciences*, *14*(12), 1212. <https://doi.org/10.3390/brainsci14121212>
- *Sadeh-Sharvit, S., Camp, T. D., Horton, S. E., Hefner, J. D., Berry, J. M., Grossman, E., & Hollon, S. D. (2023). Effects of an Artificial Intelligence Platform for Behavioral Interventions on Depression and Anxiety Symptoms: Randomized Clinical Trial. *Journal of Medical Internet Research*, *25*, e46781. <https://doi.org/10.2196/46781>
- Sesame. (27.2.2025). *Crossing the uncanny valley of conversational voice*. Haettu 20.4.2025 osoitteesta https://www.sesame.com/research/crossing_the_uncanny_valley_of_voice
- Shankar, K. K., Walker, M., Frost, D., & Orrell, M. W. (1999). The development of a valid and reliable scale for rating anxiety in dementia (RAID). *Aging & Mental Health*, *3*(1), 39-49. <https://doi.org/10.1080/13607869956424>
- Sheikh, H., Prins, C., & Schrijvers, E. (2023a). Artificial Intelligence: Definition and Background. Teoksessa *Mission AI*. Research for Policy. Springer, Cham. https://doi.org/10.1007/978-3-031-21448-6_2
- Sheikh, H., Prins, C., & Schrijvers, E. (2023b). AI Is Leaving the Lab and Entering Society. Teoksessa *Mission AI*. Research for Policy. Springer, Cham. https://doi.org/10.1007/978-3-031-21448-6_3
- Spielberger, C. D., & Reheiser, E. C. (2009). Assessment of emotions: Anxiety, anger, depression, and curiosity. *Applied Psychology: Health and Well-Being*, *1*(3), 271-302. <https://doi.org/10.1111/j.1758-0854.2009.01017.x>
- Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *STAI manual for the state-trait anxiety inventory: ("Self-evaluation questionnaire")*. Consulting Psychologists Press.
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097. <https://doi.org/10.1001/archinte.166.10.1092>
- Straw, I., & Callison-Burch, C. (2020). Artificial Intelligence in mental health and the biases of language-based models. *PLOS One*, *15*(12). <https://doi.org/10.1371/journal.pone.0240376>
- *Suharwardy, S., Ramachandran, M., Leonard, S. A., Gunaseelan, A., Lyell, D. J., Darcy, A., Robinson, A., & Judy, A. (2023). Feasibility and impact of a mental health chatbot on

postpartum mental health: A randomized controlled trial. *AJOG Global Reports*, 3(3), 100165.
<https://doi.org/10.1016/j.xagr.2023.100165>

Terveyden ja hyvinvoinnin laitos. (21.9.2023). *Kouluterveyskysely 2023: lapsia ja nuoria kuormittavat monet tekijät – selkeät arkirutiinit ja ajoissa saatu apu on tärkeää*. Haettu 7.4.2025 osoitteesta <https://thl.fi/-/kouluterveyskysely-2023-lapsia-ja-nuoria-kuormittavat-monet-tekijat-selkeat-arkirutiinit-ja-ajoissa-saatu-apu-on-tarkeaa?redirect=%252Ffi%252F>

Terveyden ja hyvinvoinnin laitos. (5.2.2025). *PROMIS-keskus*. Haettu 22.3.2025 osoitteesta <https://thl.fi/aiheet/toimintakyky/toimintakyvyn-arviointi/promis-keskus>

Thunnissen, M. R., de Jong, P. J., Weijermans, J., Vet, L. J. J., Rijkeboer, M. M., & Nauta, M. H. (2024). Negative flashforward imagery in adolescent social anxiety disorder: A pilot study of imagery reports and a short EMDR intervention. *Clinical Psychology & Psychotherapy*, 31(5), e3063. <https://doi.org/10.1002/cpp.3063>

Timmons, A. C., Duong, J. B., Simo Fiallo, N., Lee, T., Vo, H. P., Ahle, M. W., Comer, J. S., Brewer, L. C., Frazier, S. L., & Chaspari, T. (2022). A call to action on assessing and mitigating bias in artificial intelligence applications for Mental Health. *Perspectives on Psychological Science*, 18(5), 1062–1096. <https://doi.org/10.1177/17456916221134490>

Vaurio, R. (2011). Symptom Checklist-90-Revised. Teoksessa Kreutzer, J.S., DeLuca, J., Caplan, B. (toim.). *Encyclopedia of Clinical Neuropsychology*, 2447-2450. https://doi.org/10.1007/978-0-387-79948-3_2012

Vorma, H., Rotko, T., Larivaara, M., & Kosloff, A. (2020). Kansallinen mielenterveysstrategia ja itsemurhien ehkäisyohjelma vuosille 2020–2030. *Sosiaali- ja terveysministeriön julkaisuja*. <http://urn.fi/URN:ISBN:978-952-00-4139-7>

Webb, C. A., Rosso, I. M., & Rauch, S. L. (2017). Internet-based cognitive-behavioral therapy for depression: current progress and future directions. *Harvard Review of Psychiatry*, 25(3), 114–122. <https://doi.org/10.1097/HRP.0000000000000139>

Weiss, D.S., & Marmar, C.R. (1997). The Impact of Event Scale-Revised. Teoksessa Wilson, J.P. & Keane, T.M. (toim.). *Assessing Psychological Trauma and PTSD*, 169-189.

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45.

Woebot Health. (2025). *Download Woebot*. Haettu 25.3.2025 osoitteesta <https://woebothealth.com/referral/>

- Wong Sarver, N., Beidel, D.C., & Spitalnick, J.S. (2014). The feasibility and acceptability of virtual environments in the treatment of childhood social phobia. *Journal of Clinical Child and Adolescent Psychology*, 43, 63–73. <https://doi.org/10.1080/15374416.2013.843461>
- World Health Organization. (2016). *International Statistical Classification of Diseases and Related Health Problems* (10. painos). <https://icd.who.int/browse10/2016/en>
- World Health Organization. (2019). 6B40 Post traumatic stress disorder. *International Statistical Classification of Diseases and Related Health Problems* (11. painos). <https://icd.who.int/browse11/l-m/en#/http://id.who.int/icd/entity/1804127841>
- Wysa. (2024). *Wysa – Everyday Mental Health*. Haettu 25.3.2025 osoitteesta <https://www.wysa.com/>
- Xu, Z., Jain, S., & Kankanhalli, M. (2024). Hallucination is inevitable: An innate limitation of large language models. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2401.11817>
- Yhdysvaltain terveystieteiden tutkimuskeskus. (27.3.2023). PROMIS. HealthMeasures. *U.S. Department of Health and Human Services, National Institutes of Health (NIH)*. <https://www.healthmeasures.net/explore-measurement-systems/promis>
- Zhong, W., Luo, J., & Zhang, H. (2024). The therapeutic effectiveness of artificial intelligence-based chatbots in alleviation of depressive and anxiety symptoms in short-course treatments: A systematic review and meta-analysis. *Journal of Affective Disorders*, 356, 459–469. <https://doi.org/10.1016/j.jad.2024.04.057>
- Zygadło, A. (2021). A Therapeutic Dialogue Agent for Polish Language. *IEEE*, 1–5. <https://doi.org/10.1109/ACIW52867.2021.9666281>
- * Systemaattiseen kirjallisuuskatsaukseen valikoituneet artikkelit

Liitteet

Liite 1.

Tietokantakohtaiset hakulausekkeet.

PubMed:

("Generative Artificial Intelligence"[Mesh] OR "Artificial Intelligence"[Mesh] OR "Deep Learning"[Mesh] OR "Large Language Models"[Mesh] OR "artificial intelligence" OR chatbot* OR "deep learning" OR "Generative artificial intelligence" OR "Generative AI" OR GenAI OR "relational agent" OR Otis OR Wysa OR Youper OR Woebot)

AND (psychotherap* OR "psychological therapy" OR "talk therapy" OR "talking therapy" OR "group therapy" OR "Psychotherapy"[Mesh] OR "Cognitive Behavioral Therapy"[Mesh] OR "cognitive behavioral therapy" OR "cognitive behavioural therapy" OR CBT OR "cognitive behavior therapy" OR "cognitive behaviour therapy")

AND ("Anxiety Disorders"[Mesh] OR "Anxiety"[Mesh] OR "Panic Disorder"[Mesh] OR "Anxiety, Separation"[Mesh] OR "Generalized Anxiety Disorder"[Mesh] OR "Neurotic Disorders"[Mesh] OR "Phobic Disorders"[Mesh] OR anxiety OR "anxiety disorder*" OR GAD OR "generalized anxiety disorder*" OR "neurotic disorder*" OR "panic disorder*" OR "separation anxiety" OR "phobi*" OR "social anxiety")

EBSCO:n tietokannat (APA PsycInfo, CINAHL, Medline ja SocINDEX):

("artificial intelligence" OR chatbot* OR "relational agent" OR "conversational agent" OR "deep learning" OR "generative artificial intelligence" OR GenAI OR DE "Generative Artificial Intelligence" OR DE "Chatbots" OR DE "Conversational Agents" OR Otis OR Wysa OR Youper OR Woebot)

AND (psychotherap* OR DE "Intervention" OR DE "Psychotherapy" OR DE "Treatment" OR "psychological therapy" OR "talk therapy" OR "talking therapy" OR "group therapy" OR "cognitive behavioral therapy" OR "cognitive behavioural therapy" OR CBT OR DE "Cognitive Behavior Therapy" OR "cognitive behavior therapy" OR "cognitive behaviour therapy")

AND ((DE "Anxiety" OR DE "Anxiety Disorders" OR DE "Social Anxiety Disorder" OR DE "Separation Anxiety Disorder" OR DE "Separation Anxiety" OR DE "Generalized Anxiety Disorder" OR DE "Social Anxiety" OR DE "Social Phobia" OR DE "Phobias" OR anxiety OR "anxiety disorder*" OR GAD OR "generalized anxiety disorder*" OR "neurotic disorder*" OR "panic disorder" OR "phobi*" OR "social anxiety disorder"))

Liite 2.

Työn tekemiseen on hyödynnetty seuraavia apuvälineitä:

Citation Machine (2025). Chegg, Inc. Käytetty lähdeviitteiden tekemiseen koko kirjoitusprosessin ajan, 01–05/2025. <https://www.citationmachine.net/>

Rayyan. (2025). Rayyan Systems, Inc. Hyödynnetty aineiston hakuvaiheessa kaksoiskappaleiden havaitsemiseen, otsikoiden ja abstraktien lukemiseen sekä artikkelien valintaan, 02/2025.

<https://www.rayyan.ai/>