JYX JYVÄSKYLÄN YLIOPISTO UNIVERSITY OF JYVÄSKYLÄ

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Ilves, Outi; Korpi, Hilkka; Honkanen, Sari; Sjögren, Tuulikki; Aartolahti, Eeva

Title: Järjestelmällisten kirjallisuuskatsausten toteutus

Year: 2022

Version: Published version

Copyright: © 2022 Kansaneläkelaitos

Rights: In Copyright

Rights url: http://rightsstatements.org/page/InC/1.0/?language=en

Please cite the original version:

Ilves, O., Korpi, H., Honkanen, S., Sjögren, T., & Aartolahti, E. (2022). Järjestelmällisten kirjallisuuskatsausten toteutus. In O. Ilves, H. Korpi, S. Honkanen, & E. Aartolahti (Eds.), Robottien, virtuaalitodellisuuden ja lisätyn todellisuuden vaikuttavuus ja merkityksellisyys lääkinnällisessä kuntoutuksessa : järjestelmälliset kirjallisuuskatsaukset (pp. 32-37). Kansaneläkelaitos. Sosiaali- ja terveysturvan tutkimuksia, 159. http://urn.fi/URN:NBN:fi-fe2022052037517

2 Järjestelmällisten kirjallisuuskatsausten toteutus

Outi Ilves, Hilkka Korpi, Sari Honkanen, Tuulikki Sjögren ja Eeva Aartolahti

2.1 Kirjallisuushaut

Robottien ja virtuaalitodellisuuden vaikuttavuutta ja merkityksellisyyttä kuntoutuksessa selvittävät järjestelmälliset kirjallisuushaut toteutettiin Jyväskylän yliopiston liikuntatieteellisessä tiedekunnassa. Aiheittain ja tietokannoittain yksilöidyt hakustrategiat suunniteltiin yhteistyössä tutkijoiden ja Jyväskylän yliopiston informaatikon kanssa.

Järjestelmälliset kirjallisuushaut toteutettiin tietokannoista The National Library of Medicine (Ovid MEDLINE), Cumulative Index to Nursing and Allied Health Literature (Cinahl), Psychological Information Database (PsycINFO) ja Education Resources Information Center ERIC. Lisäksi aiemmin julkaistujen järjestelmällisten katsausten sekä tähän katsaukseen valikoitujen laadullisten alkuperäistutkimusten lähdeviitteitä käytiin läpi mahdollisten tietokantahakujen ulkopuolelle jääneiden tutkimusartikkeleiden löytämiseksi.

Haut määrällisiin järjestelmällisiin katsauksiin, satunnaistettuihin kontrolloituihin tutkimuksiin (*randomized controlled trials*, RCT), sekä laadullisiin alkuperäistutkimuksiin suoritettiin erikseen. Lisäksi näissä kaikissa tehtiin erilliset kuntoutusrobotteihin sekä virtuaalitodellisuuteen liittyvät haut, jolloin järjestelmällisiä katsauksia toteutettiin yhteensä kuusi.

Avainsanoina käytettiin terapioihin, kuntoutusammattilaisiin, interventioihin sekä tutkimustyyppeihin liittyviä termejä. Haut tehtiin ilman kielirajoitusta ja ilman aloituspisteen aikarajaa marraskuun 2019 ja syyskuun 2020 välillä. Eri hakujen tarkat ajankohdat sekä esimerkit hakustrategioista ja niissä käytetyt hakusanat on esitetty liitteissä 1 ja 2.

2.2 Aineiston valinta

Vaikuttavuutta selvittäneiden järjestelmällisten katsausten ja RCT-tutkimusten mukaanotto- ja poissulkukriteerit olivat samanlaiset tutkimusasetelmaa lukuun ottamatta. Vaikuttavuuskatsausten mukaanottokriteerit muodostettiin PICOs-asetelman (*Population, Intervention, Comparison, Outcome, study design*) mukaan. PICOs-lausekkeen mukaisesti potilasryhmänä (P) olivat lääkinnällistä kuntoutusta tarvitsevat lapset ja aikuiset ilman ikärajoituksia ja kuntoutusinterventiossa (I) hyödynnettiin kuntoutusrobotteja tai virtuaalitodellisuutta. Kontrolliryhmän interventiona (C) hyväksyttiin tavanomainen hoito, kuntoutusta odottava ryhmä ja erilainen kuntoutuksen toteutustapa. Lopputulosmuuttujina (O) käytettiin ICF-luokituksen mukaisia ruumiin/kehon toimintoja ja rakenteita, suorituksia ja osallistumista sekä elämänlaatua, fyysistä ja psyykkistä hyvinvointia, avuntarvetta sekä opiskelu- ja työkykyä (WHO 2002).

Tutkimukset rajattiin järjestelmällisiin katsauksiin ja satunnaistettuihin, kontrolloituihin tutkimuksiin (RCT), joita pidetään tutkimusasetelmaltaan (s) vahvimpina interventioiden tehokkuuden arvioinnissa (Chandler ym. 2019; McKenzie ym. 2019). Sekä *parallel-* että *cross-over*-asetelmaa hyödyntäneet RCT-tutkimukset hyväksyttiin mukaan. Satunnaistamattomat kontrolloidut tutkimukset, kontrolloimattomat seurantatutkimukset, tapaustutkimukset sekä laadulliset tutkimukset suljettiin pois RCT-tutkimusten katsauksessa.

Määrällisen PICOs-rajauksen perusteella rajattiin katsauksen ulkopuolelle myös tutkimukset, joissa osallistujilla ei ollut kuntoutumiseen vaadittavaa sairautta tai vammaa ja siihen liittyvää suoritus- ja osallistumisrajoitetta. Koska katsausten tarkoituksena oli selvittää kuntoutusinterventioiden vaikuttavuutta toimintakykyyn, suljettiin lisäksi katsauksesta pois yksittäisen interventiokerran välittömiä reaktioita selvittävät tutkimukset.

Aineiston valinnassa kokotekstivaiheessa tutkimusten poissulkusyinä käytettiin seuraavaa jaottelua: 1) tutkimus ei täyttänyt ennalta määriteltyä PICOs-asetelmaa, 2) kyseessä oli tutkimuksen rekisteröintidokumentti (poissulku vasta kokotekstivaiheessa), 3) kyseessä oli tutkimusprotokollan kuvaus (poissulku vasta kokotekstivaiheessa), 4) kyseessä oli duplikaatti eli kaksoiskappale ja 5) kyseessä oli konferenssiabstrakti tai -posteri ilman kokotekstiä.

Merkityksellisyyttä selvittävissä laadullisten alkuperäistutkimusten katsauksissa mukaanottokriteerit muodostettiin käyttäen laadullista PICos-strategiaa (P = *Participants*, I = *Interest*, Co = C*ontext*, s = *study design*), jossa osallistujina (P) olivat erilaiset kuntoutujat, läheiset ja laajasti eri kuntoutusalojen asiantuntijat, tutkimuksen kiin-nostuksen kohteena (I) oli kuntoutusrobottien tai virtuaalitodellisuuden käyttöön liittyvät kokemukset tai käsitykset, kontekstina (Co) monialainen tai alakohtainen kuntoutus, jossa hyödynnettiin robotteja tai virtuaalitodellisuutta ja laadullisilla tut-kimusmenetelmillä (s) tehdyt tutkimukset. Laadullisen PICos-kohderyhmärajauk-sen perusteella suljettiin pois tutkimukset, joissa osallistujat olivat terveitä tai joissa oli kyseessä erilaiset teknologian kehittämiseen liittyvät käytettävyystutkimukset il-man kuntoutustoimintaa.

Sekä määrällisten RCT-tutkimusten ja järjestelmällisten katsausten interventio (I) että laadullisten tutkimusten tutkittavana oleva ilmiö (I) määriteltiin erikseen kuntoutusrobotteja ja virtuaalitodellisuutta koskeviin hakuihin seuraavin tarkennuksin: kuntoutusrobotteina huomioitiin Hennalaa ym. (2017, 16) sekä Kyrkiä ym. (2015) mukaillen kuntoutukseen ja proteeseihin (ylä- ja alaraajan kuntoutusrobotit, proteesit, kehon ulkopuoliset tuet) sekä henkilökohtaiseen fyysiseen, kognitiiviseen ja sosiaaliseen apuun tai opetukseen (mm. syöminen, liikkuminen, esineiden nostaminen ja kantaminen, siivous, itsehoidon tuki, kumppanirobotit, vuorovaikutuksen tuki, kognitiivinen tuki) liittyvät robotit. Pois jäivät kaikki sellaiset robotit, jotka eivät kuu-

lu lääkinnällisen kuntoutuksen piiriin, esimerkiksi lääketieteelliseen hoitoon tarkoitetut leikkaussalissa toimivat robotit sekä logistiikkaan ja henkilöiden nostamiseen liittyvät laitosympäristön robotit. Virtuaalitodellisuutena interventioissa huomioitiin ne virtuaalitodellisuutta, lisättyä todellisuutta tai peliharjoittelua sisältävät tutkimukset, joissa toteutui reaaliaikainen vuorovaikutus virtuaalisen ympäristön kanssa, joissa harjoittelu jäljitteli todellista toimintaa, tilaa tai ympäristöä ja joissa virtuaalinen ympäristö mahdollisti käyttäjälleen läsnäolon kokemuksen. Pois jäivät kaikki sellaiset tutkimukset, joissa kyseessä oli videoneuvottelu tai vastaava etäkuntoutukseen liittyvä interventio, vaikka näistä olisi kyseisten tutkimusten raportoinnissa käytetty virtuaalisuus-termiä.

Kielirajoituksia tutkimuksille ei tietokantahakuvaiheessa asetettu, mutta tutkimuksissa tuli olla otsikko ja abstrakti englanniksi, saksaksi, ruotsiksi tai suomeksi. Koko julkaisun perusteella aineistoksi hyväksyttiin englannin-, ruotsin-, saksan- tai suomenkieliset tutkimukset. Mukaanotto- ja poissulkukriteereiden mukaisen tutkimusten valinnan toteutti kaksi tutkijaa itsenäisesti. Valintaan vaadittiin kahden tutkijan konsensus. Jos ilmeni eriäviä mielipiteitä, pyydettiin kolmannen tutkijan mielipide. Määrällisten PICOs- ja laadullisten PICos-kriteereiden sekä poissulkusyiden toimivuutta ja yhtenäistä ymmärrystä työryhmän jäsenten kesken tarkasteltiin jokaisen valintavaiheen alussa valintojen yhdenmukaisuuden avulla ja tarvittavista tarkennuksista tehtiin konsensus. Tutkimusten valinta toteutettiin ensin otsikon ja tiivistelmän, sitten koko julkaisun perusteella. Tietokantahakutulosten hallinnassa, duplikaattien poistossa, aineiston valinnassa sekä laadunarviointien koonnissa käytettiin Covidence-ohjelmaa (*Covidence systematic review software*).

Lähteet

Chandler J, Cumpston M, Thomas J, Higgins JPT, Deeks JJ, Clarke MJ. Chapter I: Introduction. Julkaisussa: Higgins JPT, Thomas J, Chandler J ym., toim. Cochrane handbook for systematic reviews of interventions version 6.0. Viitattu 14.2.2020.

Hennala L, Koistinen P, Kyrki V ym. Robotics in care services. A Finnish roadmap. ROSE consortium (full list of contributors at the end of document). Viitattu 4.1.2021.

Kyrki V, Coco K, Hennala L ym. Robotit ja hyvinvointipalvelujen tulevaisuus (ROSE-konsortio). Helsinki: Suomen Akatemia, Tilannekuvaraportti 2015. Viitattu 14.2.2020.

McKenzie JE, Brennan SE, Ryan RE ym. Chapter 3: Defining the criteria for including studies and how they will be grouped for the synthesis. Julkaisussa: Higgins JPT, Thomas J, Chandler J ym., toim. Cochrane handbook for systematic reviews of interventions version 6.0. Päivitetty 7/2019. Viitattu 16.2.2020.

WHO. Towards a common language for functioning, disability and health. ICF —The International Classification of Functioning, Disability and Health. Geneva: WHO, 2002.

2.3 Tutkimustiedon kriittinen arviointi ja yhteenveto

2.3.1 Kriittisen arvioinnin menetelmät

Järjestelmällisten katsausten kriittinen arviointi toteutettiin AMSTAR2-menetelmällä (*A Measurement Tool to Assess Systematic Reviews*), joka on kehitetty terveydenhuollon satunnaistettuja ja/tai satunnaistamattomia interventiotutkimuksia yhdistävien katsausten luotettavuuden arviointiin (Shea ym. 2017). AMSTAR2-arviointi koostuu 16 arviointikohteesta, joissa arvioidaan mm. katsauksessa käytettyjä hakumenetelmiä, tutkimusten valintaprosessia, kuvailua ja harhan riskin arviointia sekä yhteenvedon menetelmiä, luotettavuutta ja tulosten tulkintaa. Arviointikohteista kriittisimpiä painotetaan yleisarviossa katsauksen tulosten luotettavuudesta (korkea, kohtalainen, heikko tai erittäin heikko).

Näytönastekatsauksiin mukaan otettujen RCT-tutkimusten laatua ja harhan riskiä arvioitiin Cochrane Risk of Bias 2 -menetelmällä (RoB2) (Higgins ym. 2019). RoB2-menetelmä sisältää viisi arvioitavaa osa-aluetta, joista harhan riski voi olla peräisin: satunnaistamisprosessi, suunnitelluista interventioista poikkeaminen, puuttuvat mittaustulokset, lopputulosmuuttujien mittaaminen sekä valikoiva raportointi (Higgins ym. 2019). Harhan riskin taso (*overall risk of bias*) määritellään matalaksi niissä tutkimuksissa, joissa suurin osa harhan riskin osa-alueista on matalalla tasolla, eikä yhdelläkään osa-alueella ole korkean tason harhan riskiä. Epäselväksi harhan riski määritellään niissä tutkimuksissa, joissa on raportoitu yksityiskohdat niin niukasti, ettei niistä voi tehdä päätelmiä harhan riskin tasosta. Korkean harhan riskin tutkimuksiksi luetaan ne, joissa yksi tai useampi harhan riskin osa-alue on korkealla tasolla. Tässä tutkimuksessa harhan riski arvioitiin tarkastellen alkuperäistutkimusta sen ensisijaisen tulosmuuttujan suunnasta.

Laadullisten tutkimusten kriittinen arviointi toteutettiin Joanna Briggs -instituutin (JBI) kymmenen arviointikriteerin perusteella (Aromataris ym. 2020). JBI-arvioinnin kymmenkohtaista tarkistuslistaa käytetään laadullisten tutkimusten metodologisen laadun arvioinnissa. Kunkin kriteerin toteutuminen arvioidaan asteikolla: kyllä, ei, epäselvä (Lockwood ym. 2015). Tarkistuslistassa arvioidaan tieteenfilosofisia lähtökohtia ja metodologiaa, tutkimuskysymyksiä tai tavoitteita, aineistonkeruumenetelmää, aineiston kuvausta, tulosten tulkintaa, tutkijan kulttuurisia ja teoreettisia lähtökohtia, tutkijan vaikutusta tutkimukseen, tutkittavien äänen kuuluvuutta, eettisten periaatteiden noudattamista ja johtopäätöksiä (Lockwood ym. 2015).

Kriittisen arvioinnin järjestelmällisille katsauksille sekä määrällisille (RCT) ja laadullisille alkuperäistutkimuksille suoritti kaksi tutkijaa itsenäisesti ja lopulliseen tulokseen vaadittiin kahden tutkijan konsensus. Tarvittaessa pyydettiin kolmannen tutkijan mielipide arviointikohteesta.

2.3.2 Näytönasteen arviointi interventioiden vaikuttavuudesta

Tässä raportissa tutkimusnäyttö esitetään näytönastekatsausten avulla aiheittain. Ne tiivistävät vaikuttavuustiedon ja osoittavat näytönasteen sanallisesti. Ensisijaisesti näytönastetta arvioitiin julkaistujen katsausten ja meta-analyysien perusteella. Mukaan otetut RCT-tutkimukset ja järjestelmälliset katsaukset luokiteltiin kuntoutujaryhmän perusteella. Mikäli kuntoutujaryhmästä oli julkaistu useita järjestelmällisiä katsauksia, arvioitavaksi valittiin ne katsaukset, joiden tuloksissa robotteja tai virtuaalitodellisuutta hyödyntävän kuntoutuksen vaikuttavuus oli erotettavissa muista kuntoutusmenetelmistä.

Näytönastekatsauksiin sisällytettäväksi harkittavia järjestelmällisiä katsauksia arvioitiin kriittisesti huomioiden tietokantahakujen ajankohta, katsauksiin sisällytettyjen alkuperäistutkimusten asetelmallinen vahvuus vaikuttavuuden lausumisen kannalta, sekä katsausten luotettavuus AMSTAR2-menetelmän avulla. Jos katsauksessa oli raportoitu näytönaste GRADE-menetelmällä (*The Grading of Recommendations Assessment, Development, and Evaluation*) (Guyatt ym. 2008), hyödynnettiin sitä tässä tutkimuksessa näytönastetta lausuttaessa. Mikäli katsauksessa ei ollut käytetty GRADE-menetelmää, tarkasteltiin huolellisesti katsaukseen sisällytettyjen alkuperäistutkimusten harhan riskiä ja sen huomioimista katsauksen analyyseissä, tuloksissa ja johtopäätöksissä. Näytönaste perustuu Hoitosuositustyöryhmien käsikirjan periaatteisiin (Honkanen ym. 2016). Näytönastetta kuvaavat kirjaimet A–D: A) on epätodennäköistä, että uudet tutkimukset muuttaisivat arviota vaikutuksen suunnasta ja suuruudesta, B) uudet tutkimukset saattavat vaikuttaa arvioon vaikutuksen suuruudesta ja suunnasta, C) uudet tutkimukset todennäköisesti vaikuttavat arvioon vaikutuksen suuruudesta ja suunnasta ja D) mikä tahansa arvio vaikutuksen suunnasta ja suuruudesta on epävarma. Näytönaste määriteltiin laadukkaimpien järjestelmällisten katsausten, niihin sisällytettyjen RCT-tutkimusten lukumäärän, tutkimustulosten yhdenmukaisuuden, kliinisen merkittävyyden ja sovellettavuuden sekä harhan riskin avulla. Kuntoutusrobotteja koskevissa näytönastekatsauksissa tulosta täydennettiin uudemmilla RCT-tutkimuksilla, jotka oli julkaistu katsauksen hakuajankohdan jälkeen. Täydennystä ei tehty laajojen ja uusimpien Cochrane-katsausten kohdalla.

Näytönastekatsauksiin sisällytetyistä katsauksista ja RCT-tutkimuksista koottiin taulukoihin seuraavat tiedot: kirjoittajat, maa, julkaisuvuosi, tutkimusasetelma, kohdejoukko, otoksen koko, koe- ja kontrolliryhmien interventio, keskeiset tulosmuuttujat sekä tutkimuksessa raportoitu tulos (positiivinen tai negatiivinen vaikutus tai ei vaikutusta). Tämän vaiheen teki yksi tutkija kerrallaan ja toinen tutkija tarkasti taulukoiden tiedot.

2.3.3 Merkityksellisyyden kuvaaminen

Laadullisista alkuperäistutkimuksista koottiin seuraavat tiedot: maa, aineistonhankinta- ja -analyysimenetelmät, osallistujia kuvailevat tiedot, tutkittava ilmiö, käytetty teknologia, alkuperäistutkimuksiin liittyvät sitaatit, päätulokset ja yhteenveto. Laadullisten tutkimusten kirjallisuushakujen tulos luokiteltiin laajan aineiston takia aluksi aineistolähtöisesti: kuntoutusroboteissa käytettyjen robottien perusteella ja virtuaalitodellisuudessa kuntoutujaryhmän perusteella (Sandelowski ja Barroso 2007, 152). Luokittelun perusteella muodostettiin neljä robottiluokkaa: avustavat robotit, fyysisen harjoittelun robotit, lasten terapiaa- ja erityisopetusta tukevat robotit sekä sosiaaliset robotit. Virtuaalitodellisuutta hyödyntävässä kuntoutuksessa muodostettiin kaksi virtuaalitodellisuusluokkaa: virtuaalitoidellisuutta hyödyntävä fyysinen harjoittelu sekä psykososiaalinen harjoittelu ja edukaatio. Aineistolähtöistä luokittelua jatkettiin robotti- ja virtuaalitodellisuusluokkien sisällä. Alkuperäistutkimuksista kerättiin autenttisia sitaatteja ja päätuloksia, jotka ryhmiteltiin edelleen muodostaen yhteenveto jokaisessa luokassa erikseen (Sandelowski ja Barroso 2007, 156).

Tulososa rakentuu neljästä itsenäisestä luvusta sekä näitä seuraavasta yhteenvetoluvusta. Tulosluvuissa esitetään näytönastekatsaukset vaikuttavuudesta ja laadullisen tutkimuksen yhteenvedot, molemmat erikseen kuntoutusrobottien ja virtuaalitodellisuuden osalta.

Lähteet

Aromataris E, Munn Z, toim. JBI manual for evidence synthesis. JBI, Päivitetty 8/2020. Viitattu 14.1.2021.

Guyatt GH, Oxman AD, Vist GE ym. GRADE. An emerging consensus on rating quality of evidence and strength of recommendations. BMJ (Clinical research ed.) 2008; 336 (7650): 924–926.

Higgins JPT, Savovi J, Page MJ, Elbers RG, Sterne JAC. Chapter 8. Assessing risk of bias in a randomized trial. Julkaisussa: Higgins JPT, Thomas J, Chandler J ym., toim. Cochrane handbook for systematic reviews of interventions version 6.0. 7/2019. Viitattu 6.3.2020.

Honkanen M, Jousimaa J, Komulainen J, Kunnamo I, Sipilä R, toim. Duodecimin hoitosuositustyöryhmien käsikirja. Päivitetty 2.12.2019. Viitattu 18.1.2021.

Lockwood C, Munn Z, Porritt K. Qualitative research synthesis. Methodological guidance for systematic reviewers utilizing meta-aggregation. Int J Evid Based Healthc 2015; 13 (3): 179–187.

Sandelowski M, Barroso J. Handbook for synthesizing qualitative research. New York, NY: Springer, 2007.

Shea BJ, Reeves BC, Wells G ym. AMSTAR 2. A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017; 21 (358): j4008. DOI: 10.1136/bmj.j4008.